THE JOURNAL OF Agricultural Economics Research

United States
Department of
Agriculture

Economic
Research
Service

Articles

Book Reviews

Vol. 42, No. 2

This periodical is published quarterly. Subscription rates are: 1 year, $8; 2 years, $15; 3 years, $21. Non-U.S. orders, please add 25 percent. Send check or money order (payable to ERS-NASS) to:
ERS-NASS
P.O. Box 1608
Rockville, MD 20849-1608
or call 1-800-999-6779. This periodical is also available from the U.S. Government Printing Office—(202) 783-3238.

The Secretary of Agriculture has determined that the publication of this periodical is necessary in the transaction of public business required by law of this Department. Use of funds for printing this periodical has been approved by the Director, Office of Management and Budget.

Contents

In This Issue

Agricultural economists who at some time have experienced methodological qualms about their work will enjoy the prickly essay by Hausman and McPherson. They challenge the idea that there is a neat separation between positive and normative propositions. Moral sentiments, like optimizing behaviors, are subject to rational discourse. As the authors acknowledge, agricultural economists are rarely afforded the time to reflect deeply on the philosophical underpinnings of their inquiries. However, I feel a few minutes with the Hausman/McPherson piece will be worth your while.

Konyar and Osborn found a host of variables that affected the participation of farmers in the Conservation Reserve Program, including farmer's age, tenure status, farm size, and land values. The interesting finding that a dollar in crop production is preferred to a dollar in program payment suggests that increased enrollment in the program will come at increased marginal cost.

Recently, the press has noted the increased concentration in the poultry industry. Some of the consequences of that concentration are observed in an article by Babula, Bessler, and Schluter. They found that changes in corn prices in recent periods were more readily reflected in poultry product prices, and sustained longer, than in earlier periods. Their results appear to imply, but they do not assert, major structural changes in the poultry industry.

Gale examines the trend in farm numbers in the United States from 1960 to 1988. He separates trend from shortrun economic effects, demonstrating that changes in the relation of farm prices received and paid to land values and interest rates had the expected connection to changes in farm numbers. Models of farm structure need to include both trend and shortrun variables.

Deterministic programming models may benefit from the addition of a probability dimension. Chebyshev's inequality, a device for converting probability constraints to deterministic constraints, has been criticized for being too crude. Kim, Schaible, and Segarro demonstrate when and how Chebyshev can be used in a programming problem.

A form of regression analysis is performed by Barnard and Pfefferman to improve estimates of land values, particularly for small areas. The traditional use of regression in land value research is to explain the causes of variation of change over time. The Barnard/Pfefferman technique is intended to improve the predictors of mean land values. They applied estimates of Midwest agricultural land prices only to nonirrigated cropland. Not too surprising was the finding that the estimates of land values supplied by County Executive Directors of the Agricultural Stabilization and Conservation Service has a most positive effect on the quality of estimate. Question: should any one source of information such as ASCS officials have an overwhelming influence on market information?

Malley addresses an error correction mechanism in dynamic specification of forecasting models. He derives the ECM and then applies it to aggregate consumption expenditure in the United States. He demonstrates how to calculate elasticities of consumption with respect to income and inflation.

Book reviews lead with Teigen's critical comparison of Tweeten's 1979 and 1989 texts on farm policy. One drawback of the new book is its lack of context and history of programs affecting agriculture. Teigen adds: "Perhaps, the past is irrelevant to policy directed at removing government from agriculture...".

Other reviews include Henderson's evaluation of Shaffer's book on community development, a particularly useful document for nonacademics involved with rural development. Hutchinson reviews two books on transportation, recommending one of them despite its weakness of deregulation. Caplan and Deaton give high marks to Australia's Bureau of Agricultural and Resource Economics for the content and classy presentation of the monograph on Japanese agricultural policies.

McCloskey has argued that economics should be a branch of rhetoric—or was it the other way around? Either way progress in a discipline or a field of inquiry requires not only persuasive argument but thoughtful response. We encourage our readers to assess carefully what they see here and respond in the form of note, comment, or article. Peer review usually is limited to two or three readers, not an infallible method even under the best of circumstances. Peer review differences are not always resolved. Further-

more, peer review is private conversation, some of which should be shared with the Journal's readers. If we accept the postulates of consumer sovereignty of the products of our profession, we are bound to publish the Journal to serve the readers, the authors only incidentally, and the Journal not at all. Your comments are welcome.

Gene Wunderlich

Agricultural Economics and the Chaos of Economic Methodology

Daniel M. Hausman and Michael S. McPherson

How should agricultural economists go about their business? If they are unfortunate enough to read some economic methodology (and it is hard to avoid doing so, since without warning, it creeps into introductory paragraphs and reflective asides), they are bound to feel confused. There is something like "received wisdom" which says that positive and normative economics must be radically divorced and that economists have really nothing to add on normative questions except better facts and better positive theory. The standard view of positive economics is much like Milton Friedman's—one wants correct and useful predictions, and one is uninterested in anything else. Philosophically inclined methodologists such as Mark Blaug are more circumspect, but the received wisdom is nevertheless that economists should stick to positive economics and should look for theories that work predictively.

On top of this vague orthodoxy, and only partly conforming to it, is the insistence of high theorists (particularly in the rational expectations school) that economists must seek structural models in which the behavioral postulates are all derivable from fundamental microeconomic theory. Despite the apparent, methodologically *laissez-faire* implications of the view that only predictive success matters, Friedman and the methodologically orthodox are virtually as insistent as the high theorists that one avoid any behavioral assumptions that cannot be shown to result from individual maximization.

Donald McCloskey more radically repudiates methodology altogether. Like Blaug and others, McCloskey points out that economists do not practice what the received methodological wisdom preaches. But, he then argues for the striking thesis that there is no reason for economists to follow any methodological preaching at all. According to McCloskey, the standards of good argument in economics are whatever standards of good argument economists accept.

All this must be baffling to agricultural economists, who cannot afford to indulge in pure abstract theory or to plunge deeply into issues in the philosophy of science. Economists have little choice except to imitate the methodological practice of their teachers and colleagues. But reflective individuals, who are concerned about their goals and want their practice to contribute to those goals, cannot avoid wondering whether there is something worrisome or potentially useful in this methodological ruckus.

So what ought one to think about economic methodology? Here are a few simple remarks along with brief hints at supporting arguments.

1. Received wisdom proclaims that not only are positive and normative questions distinct, but on matters of values only grunts and clubs are called for, not civilized and rational argument. This view is absurd and destructive, because people argue about even the most emotionally charged ethical issues, and their arguments are not completely without effect. There are intellectually respectable ethical problems, and an education in economics should not disqualify one from contributing to their solution. Sound assessment of economic policy alternatives requires defensible moral foundations as well as correct positive analysis.

Some economic theorists have made important contributions to foundational questions in moral theory in recent years, but applied economists do not need to become rare philosophers to cope constructively with the ethical dimensions of their work. In fact, the tools economists use in their daily work, such as cost-benefit analysis, discounting of future returns, and assessment of externalities, are morally defensible as part of the normative apparatus of a pluralistic liberal society. Forthrightness in acknowledging the normative commitments behind policy analysis leads not to nihilism but to a more critical awareness of both the strengths and the limitations of these tools.

2. Friedman's view is generally taken to be that one should regard economic theory as a set of tools that does some jobs well. If a hypothesis works, no questions need be asked about whether its assumptions (basic propositions) are realistic (true). If a hypothesis doesn't work, scrap it. If, analogously, a computer program solves some problems correctly, one might say, "Why worry about the correctness of the algorithms upon which the program is based?" But, the analogy shows what is wrong with Friedman's view. To know that the program solves all the significant problems correctly, without assessing the algorithm and code, one would need to know the answers already, and one would not need the program. To say that a hypothesis

---

Hausman is an associate professor in the Department of Philosophy, University of Wisconsin-Madison. McPherson is Herbert H.K Lehman Professor of Economics at Williams College, Williamstown, MA. They are co-editors of the journal, *Economics and Philosophy*.

"works" means only that it has worked in some sample. This success may give good reason to believe that the hypothesis will work in other cases, but one may also want to examine the assumptions (the algorithm of the program). In that way, one may anticipate successes or failures in new applications. When hypotheses or programs fail, the most efficient way of modifying them involves examining the realism of their assumptions or the peculiarities of their algorithms or codes. So, agreement with Friedman's practical, predictive vision of economics does not imply that one should treat theories as black boxes. One should take seriously questions about the realism of assumptions and the truth of irrelevant or uninteresting predictions.

3. With their insistence on structural models, the rational expectations theorists go part of the way toward the view defended in the paragraph above, for they agree that mere predictive "success" is not necessarily a good basis for policy. They do not agree however, that one should take the assumptions of models literally and seriously and attend to their empirical weaknesses. Instead, the rational expectations theorists exalt the basic model of individual rationality and self-interest to a status little short of *a priori* necessary truth. Their work results in theories with striking elegance, system, and unity, and their efforts are justified, for there is a good deal of truth to the image of agents with complete and transitive preferences motivated by nothing except the size of their own bundle of commodities and services. But such a picture is obviously not the whole truth. Other behavioral generalizations, no matter how disconnected from or even inconsistent with the basic model, may sometimes be a better basis for prediction and explanation.

4. Economists do not behave the way that Blaug or Friedman say they ought to. But Blaug and Friedman are mistaken about how economists ought to behave, and McCloskey is wrong to argue that (apart from the norms of honest and polite conversation) there are no rules—that anything goes that in fact persuades. We cannot sort through the relevant epistemological issues in a paragraph or two, nor do agricultural economists need to master this complex area of philosophy in order to do their work. What we would instead suggest is this: (a) that agricultural economists reflect on the kinds of arguments they make and heed (as McCloskey rightly urges); (b) that agricultural economists reflect about what they aim to achieve in particular inquiries and what constraints they face; and then (c) that agricultural economists consider what sort of arguments could rationally help achieve their goals, given their constraints. If all is as it should be, the arguments that, according to step a, in fact persuade are the ones that, according to step c, ought to persuade. But all is not, of course, necessarily as it ought to be in this very much less than best of all possible worlds.

5. Economists do many different things and, except perhaps at the highest level of abstraction, the same methods of theory construction and appraisal do not apply to all activities. In thinking about what theories to consider, how to appraise and modify them, and how much credence to give them, agricultural economists need to take into account the goals, constraints, and data peculiarities that apply to their activity. (Since we are not agricultural economists and do not know the discipline in detail, we cannot formulate or address the most pressing, practical, and important methodological questions that agricultural economists face.)

Although methodological reflection could not radically and rapidly change agricultural economics, even if it needed changing, such reflection is unavoidable, potentially confusing, but conceivably valuable. Methodological reflection is unavoidable, because agricultural economists who are serious about their work must consider what they want to accomplish and what is the best way to accomplish it. Methodological reflection is most likely to become confusing when it is abstract and purely philosophical, because economists for the most part are not competent epistemologists. Such reflection is most likely to be valuable when it is undertaken in relationship to the practical problems and objectives of inquiry.

A National-Level Economic Analysis of Conservation Reserve Program Participation: A Discrete Choice Approach

Kazim Konyar and C. Tim Osborn

Abstract. *The decision to participate in a Government program can be viewed as a discrete choice problem, where a farmer will choose to sign up for the program if the expected utility of participating outweighs the expected utility of not participating. In this article, the probability of farmer participation in the Conservation Reserve Program (CRP) is modeled as a discrete choice problem and the model is estimated based on data for the entire United States. Results from the first three CRP signup periods indicate that farm tenure, farm size, land value, farmer age, erosion rate, and expected net returns with and without participation influence the probability of CRP participation. These results can be useful in evaluating how farmers might react to similar programs.*

Keywords. *Conservation Reserve Program, discrete choice approach, logistic regression, minimum chi-square method.*

The Conservation Reserve Program has been seen as an effective policy instrument for achieving environmental benefits and may be used as a model for other environmental legislation. In this paper, we develop and estimate a model of CRP participation. The analysis includes important economic variables not considered in the previous studies. In addition, unlike the previous CRP modeling efforts, the empirical analysis in this study is based on national data rather than data from a few counties. Conclusions drawn from the model are useful in assessing the importance of different factors affecting participation in an expanded CRP as well as in designing and implementing other cropland retirement programs.

The CRP is a voluntary cropland retirement program established in the Conservation Title (XII) of the Food Security Act of 1985 (PL 99-198). Its purpose is to assist owners and operators of highly erodible cropland in conserving and improving the soil and water resources of their farms and ranches, improving offsite environmental quality, and limiting the production of surplus commodities. Up to 40-45 million acres of highly erodible cropland are to be placed into the CRP by the end of 1990.

Participants in the CRP must place highly erodible cropland into grasses, trees, or other acceptable conserving uses for 10 years. They must also agree not to harvest, graze, or make other commercial use of the forage for the duration of the contract, except where the Secretary of Agriculture permits, as in a drought or similar emergency. In exchange, the U.S. Department of Agriculture (USDA) pays participating farmers annual per acre rent and one-half the cost of establishing a permanent land cover (usually grass or trees). The rental payment is determined on a per farm basis through a pseudobidding process in which the farmer indicates the amount of land to be enrolled and a yearly rental payment (rental bid). After verifying that eligibility conditions have been met, county Agricultural Stabilization and Conservation Service (ASCS) committees review the farmer's application. The application is accepted and a contract is signed if the rental bid does not exceed a predetermined maximum and if the rental bid is consistent with market rents for comparable cropland. Nearly 31 million acres had been enrolled in the CRP during the eighth signup period in February 1989.

An estimated 101 million acres of highly erodible cropland meet the physical requirements for CRP enrollment. Most of this cropland is in the Corn Belt, Northern Plains, Southern Plains, and Mountain regions. Enrollment in the CRP, however, is limited to 25 percent of the cropland in a county. The 25-percent limit reduces the amount of highly erodible cropland eligible for CRP enrollment to about 70 million acres.

Previous Studies

Several studies have analyzed CRP participation. Studies by Boggess; Ervin and Dicks; and Jagger developed theoretical models of the determinants of CRP participation *(1,4,6)*.[1] They did not, however, provide any empirical estimations. Esseks and Kraft, and Kula estimated the relationship between CRP participation and farm and farmer characteristics based on limited geographical data *(7,5)*. Kula's study was based on data from one county in Missouri. He found that a farmer's age and tenure status were more significant in explaining the probability of entering the CRP than were farm size, erosion, rate, and cash

The authors are agricultural economists with the Resources and Technology Division, ERS.

[1]Italicized numbers in parentheses cite sources listed in the References section of this article.

rents. Esseks and Kraft used data from a survey conducted in four midwestern study sites. Among their results was the negative relationship of CRP participation to income from farming and the positive relationship to income from annual crops and percentage of land with erosion problems.

The empirical analysis in this paper is based on data from the first three CRP signup periods. Unlike previous empirical analyses of CRP participation, the data used in this study cover the entire country, and the analysis incorporates additional economic variables such as land values, profitability of CRP participation, and the expected profitability of continued crop production. Previous empirical analyses were based on farm-level data. In this model, the CRP participation rates and explanatory variables are evaluated at the regional level. Therefore, statistical inference on CRP participation can be made only at the regional level, not at farm level.

Model

Farmers with eligible fields have the choice of enrolling in the CRP. USDA pays enrolled farmers annual rental for 10 years and half of the cost of establishing a vegetative cover. If farmers choose not to enroll, they presumably continue to earn income associated with crop production. By its nature, the CRP participation decision can be viewed as a dichotomous choice problem.

If a farmer's behavior is consistent with a well-defined utility function in a dichotomous choice setting, a rational farmer will compare the expected utility received during the 10-year period of participation in the CRP with the expected utility received during the same period of nonparticipation. If the expected utility of participation is greater than the expected utility of not participating, the farmer will choose to participate. The stochastic utility of participation for the i^{th} farmer who participates (superscript 1) can be written as *(2)*:

$$U_1^i = V(x_1^i, s_i) + e(x_1^i, s_i), \tag{1}$$

where $V(x_i^1, s_i)$, is a real-valued function that denotes the mean utility associated with participation; $e(x_i^1, s_i)$ is a random component in utility; x_i^1 is a vector of observed attributes associated with participating in CRP; and s_i is a vector of socioeconomic characteristics of the farmer or the farm. The utility of non-participation (superscript 0) can be expressed as:

$$U_i^0 = V(x_i^0, s_i) + e(x_i^0, s_i), \tag{2}$$

where x_i^0 is a vector of observed attributes associated with nonparticipation.

A farmer will likely join the program if:

$$U^1 > U^0. \tag{3}$$

Because the utility values are stochastic, the event that condition 3 holds will occur with some probability rather than with certainty. That is, the event that the i^{th} farmer will participate in the program (P_i^1) is determined from the condition,

$$P_i^1 = \text{Prob}\,[\,U_i^1 > U_i^0\,], \tag{4}$$

or substituting in equations 1 and 2:

$$P_i^1 = \text{Prob}\,[\,e(x_i^0, s_i) - e(x_i^1, s_i) < V(x_i^1, s_i) - V(x_i^0, s_i)]\,. \tag{5}$$

The form of the function V is based on the theory of individual choice behavior. For the purposes of this empirical analysis, V is assumed to have the following linear functional form:

$$V(x\,, s) = \sum_k Z^k(x, s) b_k = Z(k\,, s)'\,b \quad k = 1, \ldots, K, \tag{6}$$

where $Z(x\,, s)'$ is a 1 by K vector of empirical functions that are used for transforming the data, b is a vector of unknown parameters to be estimated, and K is the number of explanatory variables.

A parametric functional form for P_i^1 can be derived by specifying a parametric joint distribution function, say G, for the stochastic terms in equation 5. Commonly used distribution functions are the cumulative normal, logistic, and Cauchy. The form of the function G is influenced by concern for computational simplicity, and, in this study, the logistic function was selected for G. This produced the following binary logit probability function of participation:

$$P^1 = \frac{1}{1 + \exp[b'\,Z(x^1, s) - b'\,Z(x^0, s)]}\,, \tag{7}$$

or by inverse transformation,

$$\log \frac{P^1}{1 - P^1} = b'\,[Z(x^0, s) - Z(x^1, s]. \tag{8}$$

Here, b_k's measure the effect of each of the k independent variables on the log of the odds of participating in the CRP versus not participating.

Explanatory variables fall into two categories. The first category has variables that measure the attributes of each choice as perceived by the individual decision-maker (vectors x^1 and x^0). The second category has variables that measure the farm's and farmer's characteristics (vector s).

Attributes of Choice

An obvious element of vectors x^1 and x^0 is the net return the farmer expects from participation and nonparticipation in the CRP for the duration of the contract. If farmers participate, they can expect a return equaling the guaranteed annual rental payment from the Government for each acre enrolled in the CRP minus the costs of maintaining a conservation cover. Boggess suggests additional benefits to CRP participation. Some farmers can receive nonmonetary as well as monetary benefits from the CRP by allowing wildlife activities on suitable CRP-enrolled acres. Labor and management released from CRP acreage can earn off-farm income. Farmers with trees planted as a CRP cover can receive additional revenue if the trees are harvested and marketed after the contract expires. Costs associated with CRP participation involve establishing and maintaining vegetative cover on the enrolled acres, plus any additional costs, such as the opportunity cost of the immobile factors of production that will be idled. With these costs and benefits of CRP in mind, the present value of the net return under the standard 10-year CRP (π^1) participation period can be expressed as:

$$\pi^1 = \sum_{t=1}^{10} \frac{PNR_t}{(1+r)^t} ,$$

where PNR_t is the expected per acre net return under CRP in year t, and r is the discount rate.

The expected revenue under nonparticipation depends on the expected revenue from producing crops on the CRP-eligible land. The expected revenue is determined by the market or support price and the actual or program yield for the crops, depending on whether the farmer is participating in other Government programs. Therefore, expected net revenue of nonparticipation (CNR) is the expected revenue from crop production minus production costs. If farmers participate in support programs, the opportunity cost of land taken out of production under the Acreage Reduction Program (ARP) must also be subtracted from the expected revenue from crop production. The present value of the expected net return outside CRP (π^0) for the same 10-year period can be expressed as:

$$\pi^0 = \sum_{t=1}^{10} \frac{CNR_t}{(1+r)^t} .$$

In equation 8, then, π^0 and π^1 are the single elements of vector x^0 and x^1, respectively. The discount rate for the net present value of calculations of participation and nonparticipation net returns may be different *(4)*. The revenue from CRP is certain, therefore a risk-free discount rate can be used. Uncertainty of income from crop production may call for use of a discount rate that includes a risk component.

Attributes of Farms and Farmers

In the second category of explanatory variables (vector s), several candidates may affect the probability of CRP participation. These candidates include such farm characteristics as land value, farm size, and erosion rate, and farmer characteristics, such as age and tenure status.

The decision to participate in the CRP may be inversely related to the market value of cropland. Farmers who wish to sell land enrolled in the CRP must reimburse the Government for CRP payments received if the new owners do not keep the land in the program. So, farmers may be reluctant to participate if their land has a high market value.

The size of the farm may also be related to farmers' CRP enrollment. Larger farms have greater capital investment that may not be easily disposed of or put to alternate uses if land is retired from production under CRP *(14)*. Having large capital investment effectively increases the cost of the CRP participation.

The average annual rate of erosion on the farm may have a positive effect on CRP enrollment. Farmers who have highly erosive lands may be more sensitive to erosion problems and, therefore, more inclined to participate in the CRP. Also, CRP provides farmers with a means of retiring highly erodible cropland in order to the meet the conservation compliance provision of the Food Security Act of 1985. Conservation compliance requires farmers with highly erodible cropland to obtain an approved soil conservation plan by January 1, 1990, and to fully implement the plan by January 1, 1995. Failure to comply causes producers to lose eligibility for USDA

program benefits for their entire farming operation during the years not in compliance. Conservation compliance results in a more cost-effective CRP because farmers who are subject to compliance should be willing to accept lower rental payments for retiring their highly erodible cropland.

A farmer's age may have either a positive or a negative effect on the CRP enrollment decision. Young farmers may be more willing to experiment with new alternatives to farming and thus may be more likely to participate in the CRP. Compared with older farmers, young farmers may have more and higher paying job opportunities in the labor market should they decide to participate in the CRP and pursue alternative employment. The young farmers are likely to have a higher debt-to-asset ratio compared with older farmers and may want to cut their debt by enrolling cropland in the CRP. Older farmers, on the other hand, may have fewer risks than young farmers, and would participate in the CRP to assure a guaranteed income on enrolled land for 10 years. Upon retirement, older farmers may choose to enroll land in the CRP as a means of reducing the area farmed. Therefore, no sign was speculated for the coefficient of the age variable.

A farmer's tenure status may also have an effect on CRP enrollment *(7)*. An owner-operator has both labor and capital invested in farming. A nonoperating owner (landlord) does not have labor invested and may not have capital invested. The transaction cost of joining the CRP will be higher for the owner-operator than for the nonoperating owner, unless alternative uses for capital and labor can be found. Nonoperating owners are more likely to enroll their eligible land into the CRP than owner-operators. In equation 8, then, land value, farm size, farmer's age, erosion rate, and tenure are the elements of vector s.

Data

The model requires data on CRP-participating farmers and on nonparticipating farmers. Data on the participating farmers, such as net return under CRP and farm and farmer characteristics, are available from ASCS. However, no nationwide data exist on farmers who are eligible but not in the CRP. The model, therefore, could not be estimated using the individual farmer as the observational unit because only participating farmers were observed.

The model can be estimated, however, if the observational unit is defined as a cell consisting of a group of CRP-eligible acres for farms and farmers that share similar characteristics. The dependent variable takes on a value equal to the proportion of acres enrolled in the CRP out of the total eligible CRP acres in that cell. Each explanatory variable takes on a value equal to the mean for the group of farms and farmers in that cell. The statistical estimation is then based on the interregional variations in the CRP participation rates, and average farm and farmer characteristics. Therefore, statistical inference on CRP participation can be made only at the regional level rather than at farm level. That is, the estimated model predicts the proportion of eligible acres that will be enrolled into the CRP in a region, given the average levels of the explanatory variables in that region.

For this study, a cell consists of the farms and farmers located in each Major Land Resources Area (MLRA). USDA's Soil Conservation Service has established 156 MLRA's covering the Nation. Each MLRA comprises several counties and often crosses State boundaries. MLRA's were chosen as the unit of observation rather than counties since the data we used to establish CRP eligibility were not statistically reliable at the county level. The boundaries of the MLRA's are defined so that the soil characteristics and growing conditions are similar within an MLRA. This similarity allowed the use of the MLRA mean value for variables, such as net returns, land value, and erosion, as the representative values in a given MLRA. The use of the mean value as the representative value for tenure, farm size, and age was not as defensible. However, there is a sufficient amount of variation in all the explanatory variables across MLRA's so that estimation of the model is statistically meaningful.

Total CRP-eligible acres in each MLRA were calculated from the 1982 National Resources Inventory (NRI) data. Acres enrolled in the CRP were obtained from ASCS. From these data, the dependent variable, P_i^1, can be constructed in the following way: view the eligible acres in MLRA i as multiple observations n_i corresponding to (x_i, s_i) and assume that in m_i of those acres, the event (enrollment in the CRP) occurred. Then the empirical probability of participation $\hat{P}_i^1$ (the dependent variable) equals m_i/n_i *(8)*.

Average MLRA net returns under CRP participation, PNR_i, were calculated as the weighted average of individual net returns in each MLRA. Individual net returns under participation equal Government CRP payments minus the farmer's cost for initial cover establishment and annual cover maintenance. We assumed that participating farmers would find other uses for their farm's fixed assets or sell them at a fair market price so no capital cost is subtracted from CRP revenue. Government payments and establish-

ment cost figures were obtained from ASCS records of individual CRP contracts. Ervin provided an estimate of per acre maintenance cost estimates *(3)*.

Average net return for the nonparticipants in an MLRA, CNR_i, was calculated on the basis of net return to crops, including hay and forage used for livestock, grown in that MLRA. Crop-specific net returns were calculated using county average crop yields and State market prices of crops obtained from USDA's National Agricultural Statistics Service. ASCS furnished the crop loan rate (by State) and national deficiency payments. Detailed State crop budgets were provided by USDA's Economic Research Service (ERS). For program crops, the effective prices received by farmers in each State were calculated as the weighted average of the market and target prices. The proportion of participating and nonparticipating acres in crop programs were used as weights. Where an MLRA crossed a State boundary, crop prices and budgets for the MLRA were calculated as a weighted average of the respective States' prices and budgets. Each State's contribution to the total acreage of the crop in the MLRA was used for weighting. The net return to a crop in an MLRA was then calculated as the return to land and management. A typical net return a nonparticipant should expect in an MLRA (CNR_{ik}) is the weighted average of individual crop net returns in that MLRA. The crops included in the analysis were corn, wheat, sorghum, barley, cotton, soybeans, oats, and hay.

We calculated the land value variable as the weighted average of county land values in each MLRA. The ERS County Land Value Survey contains the annual county land value estimates. The erosion variable was calculated as the average annual soil loss (tons/acre) from sheet, rill, and wind erosion aggregated to the MLRA level from 1982 NRI data. We aggregated the remaining explanatory variables to the MLRA from the 1982 Census of Agriculture. The tenure variable was calculated as the ratio of crop acreage occupied by full tenants to total crop acreage in a given MLRA. "Full tenant" is defined as farm operators that rent all the land that they cultivate. Acreage cultivated by full tenants is identical to acreage owned and rented out by the nonoperating owners. The farm size and age variables became the average farm size and the average age of farmers in a given MLRA.

Estimation

We estimated the model by using data from the first three CRP signups held in 1986. The more recent data were not included in the analyses for two reasons. First, in the initial signups, farmers did not have the knowledge of the bid caps. Therefore, their bids were more likely to represent their true reservation price. Second, in some counties, after the third signup, the maximum acreage enrollment limits of 25 percent were being reached, and this kind of constraint could not be successfully incorporated into the model.

Several expectation formations for PNR and CNR were tried in preliminary analyses, and the naive expectation formation was selected. Thus, the expected net returns in 1986 were presumed to equal actual 1985 net returns. The expected net returns were assumed to stay constant over the 10-year CRP contract period. Assuming net returns stayed constant over the duration of enrollment meant net returns from 1 year were used instead of the present value over 10 years. This simplifying assumption produces similar statistical results and implicitly discounts both PNR and CNR at the same rate. The actual equation used for estimation took the following form:

$$\begin{aligned}\log \frac{\hat{P}^1_i}{1-\hat{P}^1_i} &= b_0 + b_1 (CNR_i - PNR_i) \\ &+ b_2\, LANDVALUE_i \\ &+ b_3\, TENURE_i + b_4\, FARMSIZE_i \\ &+ b_5\, AGE_i + b_6 EROSION_i + u_i. \qquad (9)\end{aligned}$$

Equation 9 was estimated using the "minimum logit chi-square" method. This method involves applying weighted ordinary least squares (OLS) to equation 9. The estimator is consistent and asymptotically normal *(8)*. Weights are equal to the variance of the error term which was estimated as:

$$Var(u_i) = 1 / (n_i \cdot \hat{P}_i^1 \cdot \hat{P}_i^2).$$

Results and Implications

Table 1 shows the regression results for the first three CRP signups. Estimated coefficients were significantly different from zero at the 1-percent level for (CNR_i - PNR_i), LANDVALUE, TENURE, FARM SIZE, and AGE, and at the 5-percent level for EROSION. All *a priori* expectations for the signs of coefficients were confirmed. Two statistical tests were used to determine the explanatory power of the model. The first was the pseudo-R^2 which equals ($WSSR_c$ - $WSSR_u$) / $WSSR_u$, where $WSSR_c$ and $WSSR_u$ are the weighted residual sum of squares from the constrained model (that is, all the coefficients except the constant term are set to zero) and the unconstrained model (that is, the model that is being estimated), respectively. This measure indicated that 69 percent of the variation in the dependent variable came from the model's explanatory variables. As a

Table 1—Estimates of the parameters of the CRP participation decision model

Parameter	Coefficient estimates[1]	Standard errors	Aggregate elasticity
INTERCEPT (signup 1)	1.7989	2.84620	
INTERCEPT (signup 2)	2.9380	2.84430	
INTERCEPT (signup 3)	3.3638	2.84310	
(CNR - PNR)	.0102[1]	.00218	-0.553[4] 0.322[5]
LANDVALUE	-.0007[2]	.00006	-.446
TENURE	12.0520[2]	2.11130	1.690
FARMSIZE	-.0003[2]	.00006	-.316
AGE	-.1580[2]	.05550	-6.950
EROSION	.0425[3]	.02570	.390

Pseudo -R^2 = 0.69
χ^2_8 = 223.70

[1]Coefficients measure the effect of the variable on the log of the odds of participation divided by the odds of nonparticipation.
[2]Significant at 1-percent level.
[3]Significant at 5-percent level.
[4]Elasticity of participation with respect to CNR.
[4]Elasticity of participation with respect to PNR.

second measure of the model's significance, we used a chi-square statistic with k-1 degrees of freedom (k is the number of exogenous variables in the model). The statistic was calculated as -2 $\ln(L_c/L_u)$, where L_c and L_u are the value of the likelihood function for the constrained and unconstrained models, respectively. The magnitude of this statistic suggested that the coefficients of the estimated model differ significantly from zero.

Equation 9 was estimated separately for each signup and for the first three signups combined. The coefficient estimates from different signups allowed for the detection of changes over time in farmers' response to the explanatory variables. An F-test was applied to determine if the estimated coefficients of subsequent signups were significantly different. Such a difference in coefficient estimates would imply structural changes in farmers' reactions to CRP participation between signups. The hypothesis that the slopes and intercepts were equal for all three signups was rejected at the 1-percent level. But, a further F-test showed that the structural change between signups came from changes in intercept and not in the slopes. This outcome suggests that although the farmers responded similarly to the explanatory variables in each signup, average participation during the first three signups increased due to some other factor. The most likely factor is the time it takes for farmers to learn about the CRP. Because CRP rental payments increased only marginally over the first three signup periods and the other explanatory variables presumably remained the same, the increase in enrollment rates that occurred was probably attributable to growing knowledge of the program by farmers. Examination of the magnitude of the intercepts with subsequent signups shows that the marginal learning effect decreased with time.

The negative coefficient value of $(CNR_i - PNR_i)$ suggests that as the magnitude of this variable decreases, the proportion of acres enrolled in the CRP increases. The magnitude of $(CNR_i - PNR_i)$ will decrease if net returns to CRP participation increase or net returns to growing crops on CRP-eligible lands decrease. The estimated coefficient of LANDVALUE indicates that in areas where the price of farmland is high, farmers are less likely to enroll acres into the CRP. Farmland prices in an MLRA may be relatively high due to high cropland productivity or because of alternative uses (for example, metropolitan development) for the land. The positive sign of the TENURE coefficient suggests that nonoperating landowners are more likely to enroll their eligible acres in the CRP than owner-operators. This finding confirms one of the main results of Kula *(7)*. The negative sign of the FARMSIZE coefficient indicates that the rate of participation in the CRP by large farms will be lower than the participation rate among smaller farms. There was no *a priori* expectation on the sign of the AGE variable coefficient. The estimated negative sign suggests that the older the farmer, the lower the probability of participation in the CRP. The positive sign of the EROSION variable shows that the probability of CRP participation is higher in areas where the soil is eroding at a higher rate.

Table 1 also shows the weighted aggregate elasticities of probability of CRP participation. The formulas developed by Domencich and McFadden give the i^{th} farmer's elasticity of participation, E^1, with respect to explanatory variable k that is related to choice 1 as:

$$E_i^1(1,k) = b_k Z_i^{1k}(1 - P_i^1),$$

and the elasticity of participation with respect to explanatory variable k that is related to choice 0, is expressed as:

$$E_i^1(0,k) = b_k Z_i^{0k} P_i^0.$$

These expressions clearly show two distinct elasticities for the choice-specific explanatory variables. The weighted aggregate elasticity is calculated by multiplying the individual elasticities by $n_i P_i^1 / \sum_i n_i P_i^1$ and summing over i.

A particular aggregate elasticity estimate measures the change in the percentage of eligible land enrolled

in the CRP resulting from a uniform 1-percent change in an explanatory variable across all observations. For example, a uniform 1-percent decrease in land values in an average MLRA, *ceteris paribus*, would bring about an additional 0.466-percent increase in the number of acres enrolled in CRP out of the MLRA's eligible acres.

An important elasticity estimate from a policy standpoint is the elasticity of probability of participation with respect to net returns. This elasticity can be separated into two different elasticities: one associated with net returns from CRP participation (PNR) and the other associated with returns from continued crop production (CNR). Both measure the percent change in the frequency of acres enrolled in CRP as their respective net returns change by 1 percent. Interpretation of the estimates shows that farmers are more responsive to changes in returns from crop production than to changes in returns from CRP participation. This result may seem counterintuitive because CRP participation provides a guaranteed income while returns from crop production are usually more risky. USDA commodity programs, however, essentially provide a guaranteed price floor to participating farmers for the commodities they produce. If market prices are high, farmers can reap even greater profits. In contrast, while CRP participation provides an income floor it also imposes an income ceiling because rental payments are constant over the duration of the contract. Moreover, farmers face significant penalties if they wish to terminate prematurely CRP contracts to resume crop production. Consequently, as we have defined these variables, *ceteris paribus*, the farmer would prefer a $1 increase in CNR to a $1 increase in PNR. This suggests that in the face of increasing returns to crop production, the Government has to increase CRP bid levels even faster to get more acres into the CRP under existing eligibility conditions.

Another policy-relevant elasticity is the elasticity with respect to erosion. Interpretation of the elasticity shows that a 1-percent increase in the average erosion rate results in a 0.39-percent increase in the number of acres enrolled as a percent of total eligible acres. Unlike rental payments, policymakers cannot vary erosion rates. However, program eligibility criteria can be altered to target more or less erodible cropland. The elasticity of erosion suggests that if policymakers enlarge the number of CRP-eligible acres by admitting new acres that are less erosive, they should expect lower rates of participation from the additional acres.

Conclusions

An important policy conclusion stemming from the results of the model is that farmers are more responsive to changes in returns from crop production than to changes in returns from CRP participation. To keep pace with increasing returns to crop production, the Government has to increase CRP bid levels even faster if more acres are to be brought into the CRP under existing eligibility conditions.

Some groups, including the U.S. General Accounting Office, have noted that the benefits of the CRP could have been improved by using different eligibility criteria or by employing different implementation strategies. However, many regard the CRP as an effective policy instrument for achieving environmental benefits from the retirement of targeted acreage. The CRP will likely be used as a policy model for additional environmental legislation. Legislative bills have been proposed that would increase actual land devoted to the CRP from 40-45 million acres to 60 million acres through 1992 and would create a wetlands reserve program modeled after the CRP. The national model developed and estimated here could be useful in assessing the relative significance of the factors that would affect additional CRP participation, as well as the design and implementation of future cropland retirement programs similar to the CRP.

References

1. Boggess, W. "Implementing the Conservation Reserve Provisions: Potential Risk Facing Farmers." Paper presented at the Southern Regional Project S-180, "An Economic Analysis of Risk Management Strategies for Agricultural Production Firms." Tampa, FL, Mar. 1986.

2. Domencich, T., and D. McFadden. *Urban Travel Demand: Behavioral Analysis.* Amsterdam: North-Holland Publishing Co., 1975.

3. Ervin, D. Personal communications. U.S. Dept. Agri., Econ. Res. Serv., 1988.

4. Ervin, D.E.,and M.R. Dicks. "Cropland Diversion for Conservation and Environmental Improvement: An Economic Welfare Analysis," *Land Economics*. Vol. 64(1988), pp. 257-68.

5. Esseks, J.D., and S.E. Kraft. "Why Eligible Landowners Did Not Participate in the First Signups of the Conservation Reserve Program," *Journal of Soil and Water Conservation*, forthcoming.

6. Jagger, C. "The Cost of Conservation Reserve: Predictions and Performance." Manuscript, Cornell University, Ithaca, NY, Dec. 1987.

7. Kula, O. "The Conservation Reserve Program: Determinants of Participation." Working paper, U.S. Dept. Agr., Econ. Res. Serv., Nov. 1988.

8. Maddala, G.S. *Limited Dependent and Qualitative Variables in Econometrics*. Cambridge: Cambridge University Press, 1983.

9. Reichelderfer, K.H., and W.G. Boggess. "Government Decision Making and Program Performance: The Case of the Conservation Reserve Program," *American Journal of Agricultural Economics*. Vol.70 (1988), pp. 1-11.

10. U.S. Department of Agriculture, Agricultural Stabilization and Conservation Service. Conservation Reserve Program data files, periodically since 1986.

11. _____. *Land Resources Regions and Major Land Resources Areas of the United States*. AH-296. 1981.

12. _____. *Loan Rates, Target Prices, and Deficiency Payment Rates, 1974-1988 Crops*. 1988.

13. U.S. Department of Agriculture, Economic Research Service. *County Land Value Survey*. Annual.

14. _____. *Economic Indicators of the Farm Sector: National Financial Summary, 1987*. ECIFS 7-1. Oct. 1988.

15. _____. *A User's Manual for the Irrigation Production Data System (IPDS)*. Staff Report No. 89-10, Apr. 1989.

16. U.S. Department of Agriculture, National Agricultural Statistics Service. *NASS County Data, 1972-1986*. 1988.

17. U.S. Department of Agriculture, Soil Conservation Service. *1982 National Resources Inventory*. 1984.

18. U.S. Department of Commerce, Bureau of the Census. *1982 Census of Agriculture*. 1984.

19. U.S. General Accounting Office. *Farm Programs: Conservation Reserve Program Could Be Less Costly and More Effective*. GAO/RCED-90-13, Nov. 1989.

Poultry-Related Price Transmissions and Structural Change Since the 1950's

Ronald A. Babula, David A. Bessler, and Gerald E. Schluter

Abstract. *We use a vector autoregression analysis of corn price, farm poultry price, and consumer poultry (meat) price for two periods, a mid-1950's to late 1960's period and an early 1970's to mid-1980's period. We explored the dynamic aspects of the estimated price transmission models of these two periods. Statistically significant evidence suggests a change in the dynamics of these price transmissions between the two periods. Price-increasing shocks to corn production are now more likely to be passed on to consumers more quickly as prices increase. Evidence also suggests that, since the early period, corn has become an increasingly important factor in explaining poultry prices, and substantial market change has occurred in the poultry industry at the retail level.*

Keywords. *Corn/poultry price dynamics, vector autoregression, Tiao-Box likelihood ratio tests, Kloek-Van Dijk Monte Carlo methods, industry structural change.*

When a policymaker considers alternative policies for a sector, the economic dynamics of the sector under consideration may influence the final decision. Understanding the nature of farm/nonfarm price transmissions is vital to competent policy formulation. For certain farm sector shocks, the time required for the farm price to affect the prices in the economy's other sectors (reaction times), as well as the directions, patterns, and durations of subsequent price responses, may influence the nature of the policy chosen. For instance, suppose the shock were a drought and the feed-livestock-meat agents were organized into a perfectly competitive market system of price-taking producers and consumers. Some livestock producers, faced with higher feed costs, may liquidate herds. In the short term, one would expect lower livestock and meat prices from increased slaughter but higher prices in the longer term after the short-term glut passes. An appropriate policy response to such conditions could be feed subsidies or transportation aid for moving feed into drought areas.

Some agents in the system, however, may hold sufficient market power such that distressed selling does not occur and cost increases get passed to the consumers. The appropriate policy response now differs from the first case. The policy response may instead focus on protection from abuse of market power through investigation and prohibition of price gouging. Thus, research that identifies the nature of the price transmission mechanisms provides useful economic intelligence for choosing the appropriate policy instrument.

We focus on the price transmission mechanisms linking farm corn price (PC), livestock or farm poultry price (PL), and consumer or meat price for poultry products (PM)[1] (the "corn/poultry," or PC-PL-PM, price transmission mechanism). We address characteristics of the corn/poultry price transmission: (1) the time required for a farm-level corn price shock to influence farm poultry and consumer meat prices (reaction time), (2) the direction and patterns in these responses of farm poultry and meat prices, (3) the statistically significant durations of responses in farm poultry and consumer prices to a farmgate shock in corn prices, (4) similarities and differences in price response patterns across the farm poultry and consumer poultry sectors, and (5) any evidence on whether the PC-PL-PM transmission mechanism has changed over time.

While our approach identifies the nature of these price transmissions, explaining the reasons for the nature of the identified transmissions remains for further research using models with more economic structure. We provide an initial phase for such inquiry. To facilitate such further research with more structural econometric methods, we discuss our dynamic results in relation to observed post-1950 changes in the poultry industry.

We apply vector autoregression, or VAR, econometrics throughout our study. We found that this method is appropriate because the data-oriented technique provides evidence on the dynamic properties of relationships. Static theory and more structurally oriented econometrics ignore these dynamic properties of relationships or, at best, treat them in an *ad hoc* manner.

VAR Econometrics

Under rather general conditions, an m-component vector, indexed by time t, admits an autoregressive

---

Babula is an agricultural economist with the Agricultural and Rural Economy Division (ARED), ERS. Bessler is a professor of agricultural economics, Texas A&M University, College Station. Schluter is leader, National Aggregate Analysis Section, ARED, ERS. We thank Dr. Ted Covey of the Finance and Tax Branch, ARED, ERS, for help in compiling information concerning post-1950 changes in the poultry industry.

[1]Unless otherwise stated, "farm poultry" and "consumer" (or "meat") prices are those of poultry products.

representation generally expressed as equation 1 (*3*).[2] (Note: underlined characters represent matrices or vectors.)

$$\underline{x}(t) = \text{SUM}(s{=}1,\text{inf})[\underline{b}(s)*\underline{x}(t\text{-}s)] + \underline{e}(t), \quad (1)$$

where, SUM(s=1, inf) is the summation operator for variable "s" over the range of 1.0 through infinity (inf). The b(s) are m*m matrices of autoregressive (AR) regression coefficients, and e(t) is an m-element vector of white noise residuals or innovations.[3] The white noise nature of e(t) satisfies equations 2 and 3 (*2,3,4*).

$$E(\underline{e}(t)) = \underline{O} \text{ for all t, and} \quad (2)$$

$$E(\underline{e}(t)\underline{e}(s)') = \underline{O} \text{ if t does not equal s;} \\ = \underline{S}, \text{ a positive-definite, m*m covariance matrix when t=s.} \quad (3)$$

"E" signifies the expected value operator. For applied work, equation 1's infinite lag sequence must be truncated to a number small enough to be operational but large enough for the residuals to approximate white noise (*2*, p. 112). A universally accepted method of VAR lag selection, however, does not exist (*3*). One choice used with some success is the Tiao-Box likelihood ratio test. Bessler (*3*) discusses the test's properties and suggests its use in applied problems.

Compared with more conventional, structural econometric analyses, VAR econometrics is a relatively new approach that reveals empirical regularities from time-ordered data. The approach imposes few or no *a priori* (theoretical) restrictions on data interrelationships. Rather, VAR models loosely utilize theory to suggest which variables constitute a dynamic system in equation 1. All variables in the system are initially considered endogenous, that is, each variable influences itself and all others in the system, with lags.

Estimated VAR Model, Data Sources, and Scenario Design

Equations 4, 5, and 6 provide a three-variable VAR of corn price (PC), farm poultry price (PL), and consumer or poultry meat price (PM). To allow us to investigate potential changes in price transmission, we formulate the monthly VAR model for two periods, an "early" period and a "recent" period defined below:

$$\begin{aligned} PC_t = {} & a_{c,0} + a_{c,T}*TRD \\ & + a_{c,1}*PC_{t-1} + \ldots + a_{c,k}*PC_{t-k} \\ & + a_{c,k+1}*PL_{t-1} + \ldots + a_{c,2k}*PL_{t-k} \\ & + a_{c,2k+1}*PM_{t-1} + \ldots + a_{c,3k}*PM_{t-k} + c_t \end{aligned} \quad (4)$$

$$\begin{aligned} PL_t = {} & a_{L,0} + a_{L,T}*TRD \\ & + a_{L,1}*PC_{t-1} + \ldots + a_{L,k}*PC_{t-k} \\ & + a_{L,k+1}*PL_{t-1} + \ldots + a_{L,2k}*PL_{t-k} \\ & + a_{L,2k+1}*PM_{t-1} + \ldots + a_{L,3k}*PM_{t-k} + L_t \end{aligned} \quad (5)$$

$$\begin{aligned} PM_t = {} & a_{m,0} + a_{m,T}*TRD \\ & + a_{m,1}*PC_{t-1} + \ldots + a_{m,k}*PC_{t-k} \\ & + a_{m,k+1}*PL_{t-1} + \ldots + a_{m,2k}*PL_{t-k} \\ & + a_{m,2k+1}*PM_{t-1} + \ldots + a_{m,3k}*PM_{t-k} + m_t. \end{aligned} \quad (6)$$

The variables PC, PL, and PM are defined above. All a-coefficients are regression coefficients; the c, L, and m subscripts on the a-coefficients refer to the PC, PL, and PM variables, respectively. TRD is a time trend capturing time-dependent influences not of direct concern to this study. The coefficients with a nought subscript are intercepts. The c_t, L_t, and m_t are the innovations for the PC, PL, and PM equations, respectively. We accounted for seasonal effects with 11 monthly indicator variables. Data are in natural logarithms. The k is the chosen lag number.

For this paper, we used more than 30 years of monthly price data. We used these data in natural logarithm (logged) form because the percent changes of the series are more apt to be stationary processes than the actual levels of the data.

For the corn price at or near the farmgate (PC), we use the Bureau of Labor Statistics (BLS) producer price index (PPI) (farm products index, corn no. 2 at Chicago category ("farmgate" corn price)). The PPI (farm products index, live poultry category) serves as the livestock or farm poultry price (PL). The consumer price index (CPI) (all urban consumers index, poultry category) is the consumer or meat price for poultry. Doan and Litterman's package, Regression Analysis of Time Series (RATS), generated all VAR econometric results (*7*).

We used a several-phased procedure to obtain evidence related to answering the questions about the corn/poultry price transmission. First, we estimated a monthly corn/poultry price VAR (equations 4, 5, and 6) for two periods: an "early" period from the start of 1956 to the end of 1968 (1956:1-68:12) and a "recent"

[2]Italicized numbers in parentheses cite sources listed in the References at the end of this article.

[3]Innovations in a VAR context differ from the usual economic meaning of innovation. Rather than meaning the deliberate introduction of a new economic process expected to continue for some indefinite future in the economic system, innovation in the VAR context refers to an unexpected "surprise" or shock, perhaps random, to the economic system. VAR econometrics then models the adjustment path as this shock reverberates throughout the economic system and eventually decays.

period (1973:1-85:11) (the early and recent models, respectively).[4] Choosing these periods allowed about half our sample of overall data for each period and provided the maximum numbers of observations for both VAR models. Also, the cutoff between the two periods roughly coincided with the early-1970's start of intensive poultry marketing efforts by restaurant and fast-food outlets (*12*, p. 14). Tiao and Box's (*14*) likelihood ratio tests, conducted at Lutkepohl's (*11*) suggested 1-percent significance level, suggest a one-order lag for the early model and a six-order lag for the recent model. So in equations 4 through 6, k or the lag number for each variable, is one for the early model and six for the recent model.

Second, we shocked the early and recent VAR's with a one-time increase in corn price. We analyzed the resulting impulse responses across prices within each of the two models, and then across the two models. This comparison provided reaction times, durations, and patterns in responses of farm and consumer poultry prices to corn price changes, as well as changes in these response patterns between the two periods.

Third, we used Sims' (*13*, p. 17) test of "structural change" to determine whether or not VAR coefficients have changed since the 1950's. Results not only suggest whether price structure has changed but where such change has occurred (in crop, farm poultry, and/or retail sectors).

Fourth, we analyzed forecast error variance (FEV) decompositions both across each VAR's component series and then comparatively across the early and recent models. These results suggest the strength of interrelationships among prices and how the nature of these price interrelationships have changed between the early and recent periods.

Influences of a Shock in Corn Price

The impulse response function simulates, over time, the effect of a one-time shock in one of a VAR's series on itself and on other series in the system (*3*). We shocked each VAR by a standard error (increase) of the historical innovation of farmgate corn price. We normalized the impulse responses of each variable by the standard deviation of each variable's historical innovation (hereinafter, the variable's standard error). The non-normalized impulse responses are percent changes in nonlogged indices. When normalized, these impulse responses become approximate percent changes in the standard error. Hence, if the normalized impulse of one price is larger (smaller) than another price's normalized impulse, then the second price has been more (less) "traumatized" or influenced by the initial corn price shock.

We imposed a Choleski decomposition on each VAR to orthogonalize the current innovation matrix, such that the variance/covariance matrix of the transformed current innovations is identity. We ordered the series as follows: corn price, farm poultry price, and consumer price in each VAR model. This ordering assumes that, in contemporaneous time, if a casual pattern does exist, it flows from corn price to farm poultry price to consumer price. Monte Carlo methods developed by Kloek and Van Dijk (*8*), and programmed by Doan and Litterman (*7*), generated a t-statistic for each impulse response.

Figures 1 and 2 provide responses in poultry prices from a standard error rise in corn price. Circled impulse responses are statistically different from zero at the 10-percent level of significance. We emphasize these statistically significant portions of the impulse response patterns in our analyses. Panels 1a and 1b provide the farm poultry price impulses of the early and recent models. Panels 2a and 2b provide the impulses in consumer poultry price for the early and recent models. Note that both figures present changes in, and not levels of, price indices.

Impulse Responses in Farm Poultry Price

During the early period, panel 1a suggests that a positive shock to corn price generated 6 months of declines in farm poultry price. Eventually, the early model's price declines approached zero, and then prices began to increase. Early model impulses in PL, however, were not statistically significant from zero.

This earlier period's patterns of impulse responses parallel those expected where producers are price-takers in a perfectly competitive industry. Poultry producers, having faced higher feed costs, marketed birds early. These earlier than expected marketings led to price-depressing higher slaughter ("expense-induced" slaughter). This expense-induced slaughter would give a pricing pattern similar to that observed

[4]Throughout, the number following a postyear colon refers to the month, with "1" representing January and "12" representing December. The following observations were saved as validation periods to ascertain the models' out-of-sample predictability relative to naive forecasts: 1969:1-1971:9 for the early model and 1985:12-1988:9 for the recent model. Forecasts were evaluated for both VAR's at the 1-, 2-, 6-, 12-, and 18-month horizons. Each equation (except the early PM) generated Theil-U statistics which about equaled, or were less than, unity at most of these horizons. Generally, the equations predicted well out-of-sample relative to the naive model.

Figure 1
Changes (Impulse responses) in livestock poultry price

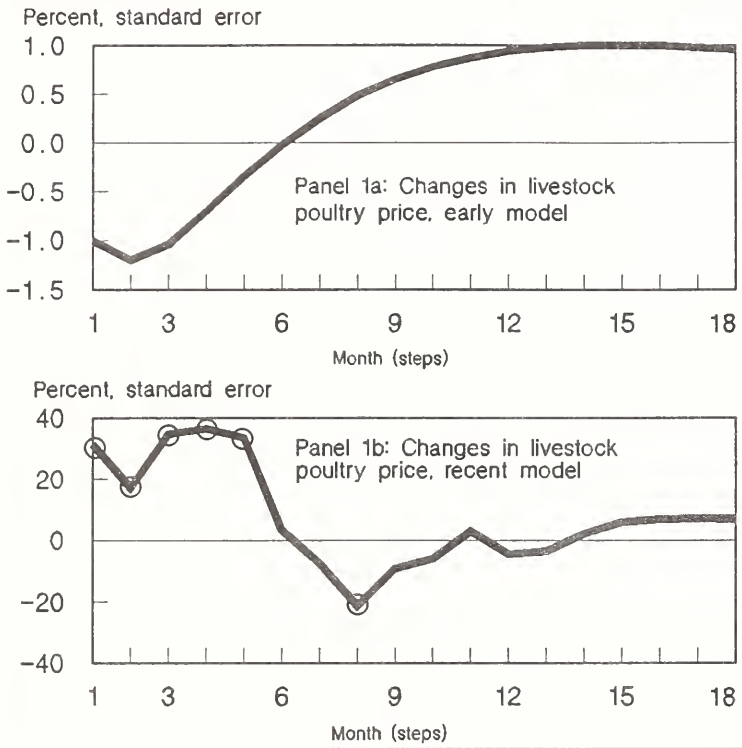
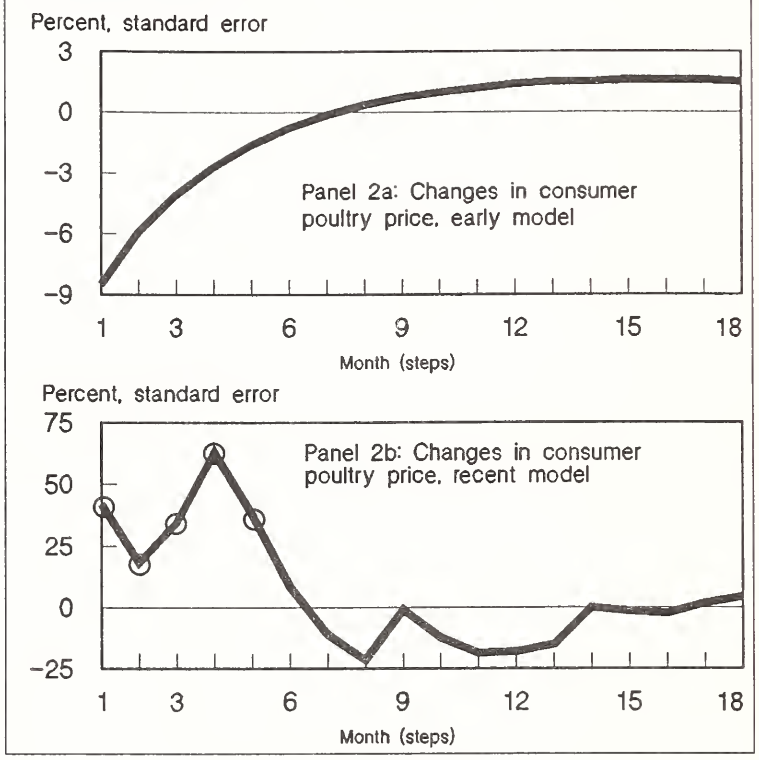


Figure 2
Changes (Impulse responses) in consumer poultry price



in the early VAR's initial half-year of farm poultry price reductions.

During the recent period, increased corn price generated farm poultry price impulses that differed from early VAR impulses. They differed in direction (increases instead of decreases) and the impulses were statistically significant. The recent VAR's farm poultry price rose in a significant manner for 5 months. (Six of the first eight impulses in the recent model's PL were significant.)

A comparison of panels 1a and 1b suggests that, in terms of PL's innovation standard error for a period, the corn price shock influenced the farm poultry price to a greater degree in the recent VAR than in the early VAR.

Two points summarize this section's results about farm poultry price impulses to a positive shock in corn price. First, the effects of corn prices on farm poultry prices drastically changed between the early and recent periods. These responses in farm poultry price changed from 6 months of statistically insignificant declines to about a half year of mostly significant increases, a change in response patterns consistent with a change from an industry of many small, price-taking producers to an industry where producers had the market power to pass on cost increases. And second, corn price shocks had far less influence on PL in the early period than in the recent period.[5]

Consumer Meat Price Responses

Consumer price impulses from the initial rise in PC closely mirrored farm poultry price response in each of the two models (figs. 1 and 2). As with farm poultry prices, consumer poultry prices took directionally opposite response patterns in the early and recent VAR's. The early model's consumer price fell in an insignificant manner for 7 months after the initial increase in farmgate corn price. Presumably, expense-induced slaughter's effects on PL filtered through to prices of retail poultry products. Yet response patterns reversed in the recent period, with consumer price having increased for about 5 months in a largely significant manner. Corn price increases influenced the recent period's consumer poultry price to a far greater extent than the early period's consumer poultry prices.

[5] A price's impulse responses are normalized by the standard error of that price's historical innovation for the model's sample period, and such innovation standard errors in the recent model exceed those of the early model for all modeled prices. Comparisons of absolute values of non-normalized impulses also suggest that recent model impulses in PL and PM exceed those of the early model.

Formal Evidence of Structural Changes

We tested coefficient constancy or "structural change" since the 1950's with the procedure of Sims (*13*, p. 17). We specified a "reduced" model for PC, PL, and PM, and estimated it over the 1954:1-1985:11 period with ordinary least squares. This "reduced" model differs from a "full model," since the full model includes test variables which were not in the reduced model. These variables are: IND; IND*PC(1), . . . ,IND*PC(6); IND*PL(1), . . . , IND*PL(6); IND*PM(1), . . . , IND*PM(6). IND is an indicator variable valued at unity after December 1970 and at zero otherwise. The asterisk denotes the multiplication operator. The parenthetical number denotes lags. A "full-vs.-reduced model" F-test was then implemented for the test variables. The null hypothesis is that the test variable coefficients are, as a group, zero, thereby implying no structural change. One rejects the null of no structural change for F-values greater than the tabular $F(19,333)$ of 1.58 for the chosen 5-percent significance level.

Evidence was insufficient to suggest that coefficients changed in the corn price and the farm poultry price relations. The F-values (1.09 for PC and 1.25 for PL) were less than 1.58. With a 2.79 F-value, evidence was sufficient to suggest structural change at the retail (PM) level. These results indicate that at least part of the cause for the directionally opposite patterns in the two model's price responses may have risen from retail-level changes.

Decompositions of FEV

Analysis of decompositions of FEV identifies the interrelationships within a modeled system's time series (*13*). Error decompositions attribute within-sample error variance to alternative series, giving a measure useful in applied work. We calculated FEV decompositions for k-step-ahead forecasts for the early and recent models (table 1).

Corn price was largely exogenous in both VAR's. The percentage of PC's FEV attributed to its own error exceeded 73 percent at all reported horizons.

One difference among the past and recent patterns of FEV decompositions was particularly evident. Farm poultry and meat prices had negligible FEV proportions attributed to farmgate corn price in the early period but much larger proportions so attributed in the recent period. The influence of corn price on poultry prices has increased between the two periods. This evidence underscores the greater statistical

Table 1—Proportions of FEV k months ahead allocated to innovations in respective series, early and recent composite model estimations

Variable name	k	Standard error	Percent explanation from PC	PL	PM
"Early" VAR estimated over 1956:1-68:12:					
Corn price (PC)	1	0.0506	99.52	0	0.48
	6	.0707	93.02	1.34	5.64
	12	.0777	83.97	5.03	11.00
	18	.0809	78.36	7.77	13.87
	24	.0825	75.57	9.22	15.21
	35	.0835	73.75	10.20	16.05
	36	.0835	73.68	10.24	16.08
Farm poultry price (PL)	1	.0576	.02	97.99	1.99
	6	.0750	.02	85.55	14.44
	12	.0820	.03	80.09	19.88
	18	.0848	.04	78.23	21.73
	24	.0860	.05	77.49	22.46
	35	.0867	.06	77.06	22.88
	36	.0867	.06	77.04	22.90
Consumer meat price (PM)	1	.0277	.59	23.38	76.03
	6	.0434	.30	37.53	62.18
	12	.0498	.24	40.93	58.83
	18	.0523	.24	41.88	57.89
	24	.0533	.25	42.22	57.54
	35	.0539	.25	42.40	57.34
	36	.0539	.25	42.41	57.34
"Recent" VAR estimated over 1973:1-85:11:					
Corn price (PC)	1	.0834	99.54	.34	.13
	6	.1392	97.97	.90	1.14
	12	.1536	92.57	1.17	6.27
	18	.1591	87.33	1.54	11.13
	24	.1607	85.65	1.53	12.81
	35	.1612	85.10	1.54	13.36
	36	.1612	85.09	1.54	13.37
Farm poultry price (PL)	1	.0751	8.60	90.59	.81
	6	.1009	18.58	73.97	7.45
	12	.1070	18.69	69.75	11.55
	18	.1077	19.13	68.69	12.17
	24	.1079	19.28	68.57	12.15
	35	.1080	19.28	68.57	12.15
	36	.1080	19.28	68.57	12.15
Consumer meat price (PM)	1	.0386	12.61	46.01	41.38
	6	.0579	23.76	54.31	21.92
	12	.0602	25.69	53.05	21.27
	18	.0607	25.54	53.18	21.28
	24	.0608	25.66	53.10	21.24
	35	.0608	25.66	53.09	21.25
	36	.0608	25.66	53.09	21.25

significance of the recent model's PL- and PM-impulse responses.

Both the early and recent VAR's suggest a minor feedback relationship from consumer price to farm poultry price. The proportion of the farm poultry price's FEV attributed to consumer meat price reached 23 percent in the early model and 12 percent in the recent model.

Consumer poultry price became more endogenous between the early and recent periods. In the early model, meat price was largely exogenous, with more than 57 percent of FEV having been self-attributed at all reported horizons. In the recent VAR, the meat price was largely endogenous with no more than about 22 percent of its FEV being self-attributed at horizons beyond 1 month.

A Collation of Results With Post-1950 Trends in the Poultry Industry

Our results concerning impulse responses, change in structure, and FEV decompositions strongly suggest that something substantial has changed in how a corn price shock pulsates through the poultry-related noncrop economy. Beginning with an increase in corn price, farm and retail poultry prices declined in a statistically insignificant manner for about 6 months in the early model. In the recent model, under the same conditions, both poultry prices not only rose for roughly 6 months, but most of these increases were statistically significant. Since the early period, perhaps producers have developed an ability to pass on corn-based feed cost increases to consumers, an ability whose existence was not evident in the early period's data. The change-of-structure test results suggest changes at the consumer or retail demand level. Patterns of FEV decompositions suggest that a more direct link has developed (since the early period) between corn price and the farm and retail poultry prices. This result underscores the higher levels of statistical significance in the impulses of both poultry prices to a corn price shock with the recent model than with the early model. The FEV decompositions suggest that retail price is more endogenous to the modeled system in the recent rather than in the early period.

Our data-oriented model has revealed that substantial and statistically significant changes have occurred, but the nontheoretical VAR models do not determine why such changes evolved. We can offer some potential reasons based on several patterns observed since the 1950's and concerning change in market structure on both the demand and supply sides of the U.S. poultry industry. Most of the patterns we discuss are for broiler and turkey prices. These two are the primary elements comprising the composite poultry prices at the producer and retail levels.[6]

Many demand- and supply-side changes have occurred in the poultry industry since the 1950's. These demand- and supply-side events are potential reasons for the changes we found in poultry price response patterns, in the dynamic characteristics of corn/poultry price interrelationships, and in the changing structure of the retail price equation. On the supply side, observed technological advancements have shifted poultry supply rightward. Demand-side changes have also occurred in the forms of both demand shifts and movements along these shifting relations. Changing poultry demand and changing broiler supply curves have interacted to generate newer and different equilibria since the 1950's. Our VAR models cannot discern the exact roles which such events have played in the data-embedded changes. We leave this task to more conventional econometric efforts which specifically address theoretical market structure issues. We have accomplished the first step in this overall inquiry by having uncovered evidence that points to sufficient changes to warrant such inquiry.

Observed Supply-Oriented Events

Lasley (*9*), the National Broiler Council (*12*), and Coffin, Romain, and Douglas (*6*) mention several post-1950 changes in the supply side of the poultry industry. These events are advancements in poultry production technology, increased flexibility of poultry supply through broiler product specialization, and more standardized poultry output quality from more vertically integrated poultry production.

Lasley (*9*, pp. 7-14) notes that over the last several decades, technologically induced production expansion has occurred so rapidly that producers responded swiftly enough to shortrun profits that prices, in the long run, have hovered very close to production costs.

[6] The composite poultry price at the farm level is the BLS producer price index (PPI) (farm products index, live poultry category). This series is comprised of (1) the PPI (farm products index, broilers category) and (2) PPI (farm products index, turkeys category). The broiler price index has an 81-percent weight and the turkey price index has a 19-percent weight in the composite poultry price at the farmgate. The CPI (all urban consumers index, poultry category) serves as the composite poultry price at the consumer or retail level. This composite consumer price is composed of (1) the CPI (all urban consumers index, fresh whole chickens category), (2) the CPI (all urban consumers index, fresh and frozen chickens category), and (3) CPI (all urban consumers index, "other" poultry category). Weights for the composite consumer price are a combined 81 percent for items 1 and 2, and 19 percent for item 3.

These technological advancements are evident from lower input requirements for poultry production. A ton of feed now produces 37 percent more broilers and 54 percent more turkey than in 1955 (*9*, p. 14). Labor input requirements per pound of broiler and turkey production are 2.6 percent and 3.8 percent, respectively, of 1945-49 requirements (*9*, p. 14).

The second supply-side event is a more flexible supply through producer/processor specialization of poultry output (*9*, *12*). This producer/processor specialization was motivated, in part, because of the various demand changes discussed below. Responding to consumer preference changes, since the early period, poultry supply today is composed of less of the traditionally demanded whole-bird product, and of more of the specialized or fully processed broiler products (*5,9,12*). More fully processed products include prepackaged birds (whole, cut-up), packaged parts, and "further-processed" products (precooked, "microwavable," and "cold cut" products) (*5,9,12*). Since 1965, proportions of federally inspected slaughter slated for cut-up products rose from 19.3 to 42.0 percent for broilers, and from 7.3 to 27.0 percent for turkeys (*9*, p. 17). Since 1960, proportions of federally inspected poultry slaughter for further processing rose from 2.7 to 10.5 percent for broilers and from 5.0 to 36.6 percent for turkeys (*9*, p. 17). The proportions of processors' marketed broilers sold as prepackaged chilled parts rose from 3.4 to 16.9 percent during 1967-87 (*12*, p. 14). These advancements have more than offset the demand-increasing events discussed below, such that by the early 1980's, deflated retail prices fell short of 1955 levels by 61 percent for broilers and by 56 percent for turkeys (*9*, p. 4).

As the third supply event, poultry output has become more standardized because of the increased vertical integration of poultry production (*6,9,12*). Since the mid-1950's, proportions of production that were vertically integrated rose from 90 to 99 percent for broilers and from 36 to 90 percent for turkeys (*9*, p. 8). The increasing concentration into fewer and larger poultry "producer/processor" firms may provide some explanation into why PL and PM responses to corn price movements are more statistically significant in the recent than in the early period. The increasing concentration may also explain why corn price uncertainty's influence on poultry price FEV's has increased since the early period. With larger diversified units came more standard blended rations. With the needs for large volumes of energy feed for these rations and the wide availability of corn as a feed energy source, it is natural that large volumes of corn would be used in these large production units. The more systematic use of purchased corn-based feed, where producers exercise more complete and more precise control over the poultry ration, may account for corn price's more direct influences on PL and PM in the recent rather than in the early model.

Observed Demand-Side Events and Interaction With a Changing Supply

Several demand-side events have together resulted in different (often augmented) consumption patterns for poultry products. Some events have shifted demand curves rightward so that more poultry is demanded at each price; other events have generated actual changes in the nature of poultry demand. Combinations in demand-supply interactions have triggered events which have caused movements along present (and previous) poultry demand curves. Since the 1950's, these changes have coincided with strong rises in annual per capita consumption: from 14 to 49 pounds of broilers and from 4.1 to 10.7 pounds of turkeys (*9*, p. 1). Either increases or changes in retail broiler consumption or both coincided with the development of consumer allegiances to brand names, ever-increasing poultry volumes marketed by restaurant and fast-food outlets, and more serious health concerns about beef and pork products than about poultry products.

Consumer allegiance to poultry brand names has been increasing, and processor marketing efforts under these poultry brands have increased consumer demand. The National Broiler Council (*12*) provides data about trends in consumer allegiance since 1981. These data are for broilers and chicken which account for 81 percent of the PL and PM indices. Today, 84 percent of poultry processors market broilers under a brand name, up from 71 percent in 1981 (*12*, p. 13). In 1987, 52 percent of marketed broilers were sold under a brand name, a proportion which rose from 39 percent in 1983 (*12*, p. 13).

Another source of enhanced consumer demand for poultry arose from heightened marketing efforts by restaurants and fast-food outlets. The share of processors' broilers shipped to restaurants dramatically increased from 1.5 percent in 1960 to 13.8 percent in 1987 (*12*, p. 14).

The net effect of growing consumer demand for poultry and the dominance of corn as an energy source in poultry rations has led to broiler rations accounting for a larger share of domestic corn usage. For example, from the early 1950's until the mid-1980's, broilers'

percentage of total feed use of corn rose from 3.5 percent to over 11.5 percent (*15*).

Findings and Conclusions

Which of the two views of the underlying market mechanisms discussed in the introduction correctly describes the effects of a price-influencing crop sector shock on the rest of the economy? Our second view did for the PC-PL-PM price transmission mechanism and for the recent period.

But, we found that our first view would have been correct in an earlier time. We found higher corn prices are passed on as higher farm poultry and consumer poultry meat prices. Under these circumstances, a public policy of feed subsidies or feed transportation aid would benefit poultry producers without protecting consumers. Thus, an understanding of the noncrop price ramifications of a crop sector shock aids policymakers attempting to formulate policy to alleviate detrimental effects of, or to cope with, a shock to a crop sector.

Our results focused on how a corn price shock pulsates through the noncrop economy, in particular for poultry-related products. We investigated how a corn price increase influences farm poultry and retail meat prices, and how such price dynamics have changed over time.

We found that if trends observed since the early 1970's continue, one may expect a rise in corn price to generate about 6 months of farm poultry and retail poultry price increases. Reaction times for PL and PM responses are immediate. Retail poultry price impulses directionally mirror impulses at the farm poultry level. Corn price appears highly exogenous, while retail meat price appears highly endogenous. The farm poultry price greatly influences the retail price. We obtained these current results about the PC-PL-PM price transmission from the recent model.

Another set of results concern how the PC-PL-PM price transmissions for poultry and broilers have changed since the early 1950's. This research reveals that marked and statistically significant changes have occurred among corn and poultry price relationships over the last 30 years. Poultry price responses (at the farm and retail levels) to rises in corn price have completely reversed. We observed about 6 months of insignificant decreases in the early period. In the recent period, this has changed to about half a year of significant increases.

Results may suggest that in the recent period the larger, more vertically integrated, and "factory-like" poultry producing/processing concerns are now somehow able to pass on rises in corn-based feed costs to consumers. This ability apparently did not exist for the more numerous and smaller poultry producers of the early period.

During the recent period, we found more statistical significance of relationships among corn, farm poultry, and retail poultry prices than in the earlier period. The analyses of the decompositions of forecast error variance of the two models underscore this finding of more statistical significance in the relationships. We conducted the Sims test for structural change on the PC-PL-PM price transmission. Evidence was sufficient at the 5-percent significance level to suggest structural change in the retail price equation.

We found a drastic change in the past 30 years in some of the dynamic characteristics of price transmission within the corn-farm poultry-poultry meat complex. Our procedures did not allow us to explain the reasons for this change. Knowledge of the change could influence the recommendations of policymakers. Explanation of the change could lead to interesting professional exchanges.

References

1. Babula, Ronald A., and David A. Bessler. "Higher Corn Prices and Livestock, Meat Prices," *Agricultural Outlook*. U.S. Dept. Agr., Econ Res. Serv., AO-147, Nov. 1988, pp. 24-25.

2. Bessler, David A. "An Analysis of Dynamic Economic Relationships: An Application to the U.S. Hog Market," *Canadian Journal of Agricultural Economics*. Vol. 32, No. 1, Mar. 1988, pp. 109-24.

3. _____. "Relative Prices and Money: A Vector Autoregression on Brazilian Data," *American Journal of Agricultural Economics*. Vol. 66, No. 1, Feb. 1984, pp. 25-30.

4. Bessler, David A., and John L. Kling. "Forecasting Vector Autoregressions with Bayesian Priors," *American Journal of Agricultural Economics*. Vol. 68, No. 1, Feb. 1986, pp. 144-51.

5. Christensen, Lee, and Bob Bishop. "Turkey Tastes Good Throughout the Year," *Agricultural Outlook*.

U.S. Dept. Agr., Econ. Res. Serv., AO-147, Nov. 1988, pp. 12-13.

6. Coffin, H. Garth, Robert F. Romain, and Meghann Douglas. *Performance of the Canadian Poultry System Under Supply Management.* Joint publication of the Department of Agricultural Economics, MacDonald College, McGill Univ., Ste. Anne de Bellevue, Quebec and le Departement d'Economie Rurale, Universite Laval, Quebec, Jan. 1989.

7. Doan, Thomas A., and Robert B. Litterman. *Regression Analysis of Time Series. Version 2.12.* Minneapolis: VAR Econometrics, 1986.

8. Kloek, T., and H.K. Van Dijk. "Bayesian Estimates of Equation System Parameters: An Integration by Monte Carlo," *Econometrica.* Vol. 46, No. 1, Jan. 1978, pp. 1-20.

9. Lasley, Floyd A. *The U.S. Poultry Industry: Changing Economics and Structure.* AER-502, U.S. Dept. Agr., Econ. Res. Serv., July 1983.

10. Litterman, R.B. "Specifying Vector Autoregressions for Macroeconomic Forecasting," *Bayesian Inference and Decision Techniques with Applications: Essays in Honor of Bruno deFinetti.* (ed. Prem Goel and Arnold Zellner.) Amsterdam: North-Holland Publishing Co., forthcoming.

11. Lutkepohl, H. "Comparison of Criteria for Estimating the Order of a Vector Autoregression Process," *Journal of Time Series Analysis.* Vol. 6, No. 1, 1985, pp. 35-52.

12. National Broiler Council. "NBC Survey Results: How Broilers are Marketed. Selected results from the National Broiler Council's Marketing Practices Survey," *Broiler Industry.* Vol. 51, No. 2, Dec. 1988, pp. 12-18.

13. Sims, Christopher. "Macroeconomics and Reality," *Econometrica.* Vol. 48, No. 1, Jan. 1980, pp. 1-48.

14. Tiao, G., and G.E.P. Box. "Modeling Multiple Time Series: With Applications," *Journal of the American Statistical Association.* Vol. 76, 1981, pp. 802-16.

15. Van Meir, Lawrence. Head, Coarse Grains Analysis Section, Crops Branch, Economic Research Service, personal communication, July 17, 1989.

How Economic Conditions Changed the Number of U.S. Farms, 1960-88

H. Frederick Gale, Jr.

Abstract. *The annual net decline in the number of farms is explained by using the lagged number of farms to account for the longrun trend and several variables representing economic conditions. The trend provides most of the explanatory power during 1960-74, while prices, land values, and interest rates explain deviations from the trend during 1975-88. Projections of farm numbers need to take into account both longrun trends and shortrun variability in economic conditions to produce meaningful estimates.*

Keywords. *Number of farms, prices, interest rates, land values, time series.*

What role is played by economic conditions in influencing changes in the number of farms? Has this role taken on increased importance in recent years as the farm economy has become more integrated into nonfarm and world economies? This article addresses these questions by estimating a time series regression model to explain changes in the number of farms during 1960-88. I used a trend component and several variables which represent economic conditions in the farm economy.

A number of recent studies has analyzed changes in the distribution of farms in different size classes (*2,5,6,10*).[1] These studies have given little attention to the effects of economic variables, while using census data from 2 or 3 census years and focusing on the transition of farms among different size classes, as well as farm entry and exit.

While these studies identify structural trends and examine hypotheses regarding firm size and growth, they fail to consider the effects of economic conditions which affect the relative profitability of farming. The studies are primarily cross-sectional, containing insufficient variation in many price and policy variables to discern statistically their effects on the number of farms.

The present study complements these more detailed studies by using annual observations over a 28-year period to estimate a regression model that explains changes in farm numbers. The data are less detailed and may be less reliable than the census data (see appendix), but they permit consideration of the effects of year-to-year variations in shortrun economic variables on the number of farms. I am less concerned with considering in detail the longrun structural forces focused on by previous studies than the attempt to identify the role played by shortrun economic conditions.

Regression Model of Changes in Farm Numbers

In any year t, the number of farms is determined jointly by longrun structural forces and by shortrun influences appearing as deviations from the longrun trend. The change in the number of farms between year t and t-1 is expressed as a first-order difference equation to represent the longrun trend plus a deviation, D_t:

$$F_t - F_{t-1} = A(F_{t-1} - F_n) + D_t, \tag{1}$$

where F_t is the total number of farms in year t, and α and F_n are constant parameters with the restrictions, $-1<\alpha<0$, $F_n>0$. The trend component with the restrictions given above states that the number of farms will decline each year by some fixed proportion α of the difference between the number of farms in the previous year and some fixed constant F_n.[2] This implies a continuous decrease in farm numbers at a declining rate which converges toward F_n from above. This model is chosen simply because it fits the data well; a theoretical model of the longrun decline in farm numbers is not attempted here. This longrun trend could result either from some dynamic adjustment process, a longrun trend in relative prices, or a change in technology.

Economic theory suggests that net entry to a competitive industry like farming will be influenced by the expected profitability of farming relative to prospective earnings in other activities. The profitability of farming is, in turn, affected by changes in economic conditions. The deviation from the longrun trend in farm numbers, D_t, is therefore expected to be

Gale is an agricultural economist with the Agriculture and Rural Economy Division, ERS. The author acknowledges the helpful comments of Hazen Gale, David Henderson, Bill Lin, and Donn Reimund.

[1] Italicized numbers in parentheses cite sources listed in the References section at the end of this article.

[2] The observed percent change in the number of farms, $\delta_t = (F_t - F_{t-1})/F_{t-1}$, is equal to $\alpha - \alpha(F_n/F_{t-1})$. Thus, a α is not identical to δ_t unless $F_n=0$, in fact $|\alpha| > |\delta_t|$, and δ_t is not constant over time, approaching zero as F_{t-1} approaches F_n.

influenced by variables representing economic conditions, X_t, expressed as a linear function of these variables:

$$D_t = b_o + \sum_{j=1}^{K} b_j X_{jt} + e_t, \qquad (2)$$

where the b_j are parameters to be estimated, and e_t is a random disturbance.

Substituting equation 2 into equation 1, expressing the X_{jt} as deviations from their means,[3] and rearranging slightly, equation 3 is obtained:

$$F_t - F_{t-1} = (-\alpha F_n) + (\alpha F_{t-1}) + \sum_{j=1}^{K} b_j X_{jt} + e_t. \qquad (3)$$

Equation 3 expresses the change in the number of farms as a linear function of the lagged number of farms and K economic variables with intercept $(-\alpha F_n)$.

This equation can be estimated with regression analysis techniques. The estimates of the coefficients b_j are estimates of the effects of the economic variables on changes in the number of farms. A positive b_j will be found for variables associated with increased net entry of farms,[4] and negative b_j will be found when a variable is associated with less net entry. The variables actually used are discussed in the following section. Failure to reject the joint hypothesis that:

$$b_1 = \ldots = b_k = 0, \qquad (4)$$

suggests that the change in the number of farms may be explained using only the longrun trend.

The estimated coefficient on the lagged number of farms yields an estimate of α, and, given the estimate of α, F_n can be obtained by dividing the intercept by $-\alpha$. Note the special case where the coefficient on the lagged number of farms is equal to zero and the intercept is nonzero. This would suggest a linear trend in farm numbers and, consequently, no tendency for the decline in farm numbers to slow down.

[3]This forces the intercept b_0 to equal zero, which facilitates the recovery of the parameter F_n from the intercept in equation 3.

[4]Net entry of farms, F_t-F_{t-1}, has been negative for most of the past 50 years. An increase in net entry means that the negative number will move closer to zero. This is a decrease in net exit.

Data on the Number of Farms

The model discussed above is applied to the official USDA estimates of the number of farms, which peaked during the 1930's and fell continuously until the end of the 1970's when 2 years of very small increases were recorded (fig. 1) (see appendix). This study considers the years 1960-88, during which the extremely rapid decline that began in the 1950's appears to have gradually slowed down, with the number of farms approaching approximately 2 million by the end of the period. This decline appears to represent a completion of the structural shifts of the 1950's which showed little sign of slowing down as labor migrated out of agriculture and average farm size increased. The period ends with the farm boom-and-bust years of the 1970's and 1980's, when the decline in farm numbers slowed to zero and became briefly positive before returning to rapid decline during the early 1980's.

The rate of change in farm numbers declined steadily in absolute value from 1960 until about 1974 (fig. 2). After this time, the rate of change in farm numbers did not follow a predictable pattern as it had during the 1960's and early 1970's, but rather fluctuated around a mean of about minus 1 percent. The peak in figure 2 in 1979-80 corresponds to the farm "boom" years of the late 1970's when exports and land values were increasing, real interest rates were low, and farming was believed to be an attractive investment. The trough coincides with the "bust" years of declining exports and land values and high real interest rates in the 1980's. This pattern leads the author to consider whether economic conditions have exerted greater influence during recent years than in earlier years when the longrun trend appears to have been the dominant influence.

Specification of Explanatory Variables

The regression model includes variables that influence the relative profitability of farming and which fluctuate from year to year, in addition to the lagged number of farms which is intended to capture the longrun trend. The variables include prices of farm output and inputs, land values, interest rates, and the ratio of nonfarm wages to farm income.

Prices are believed to play a key role in determining entry to and exit from a competitive industry. The ratio of the index of prices received by farmers to the index of prices paid by farmers represents the ratio of

Figure 1
Number of U.S. farms, 1910-88

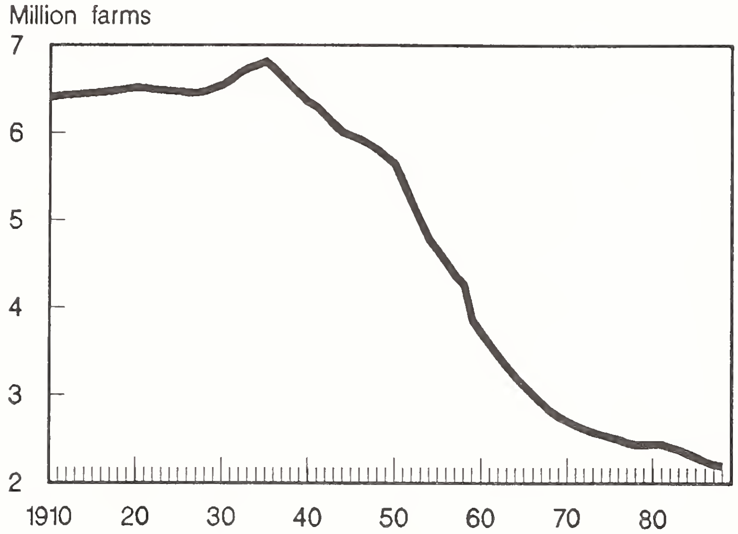
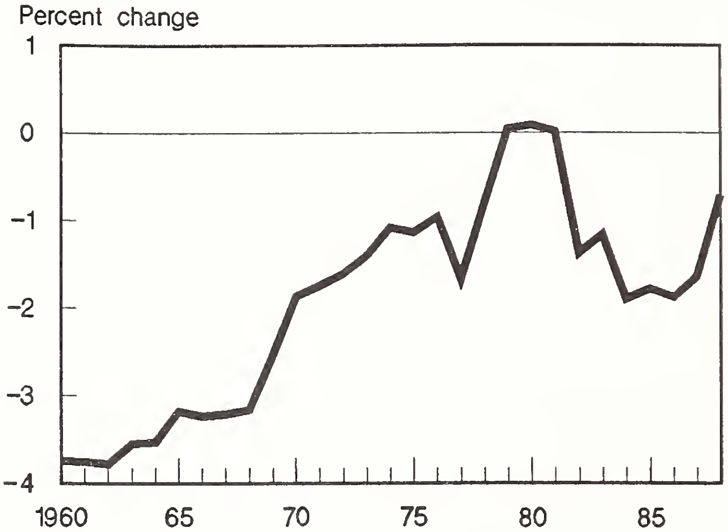


Figure 2
Rate of change in farm numbers, 1960-88



output prices to input prices. A higher value of this ratio is expected to lead to greater net entry of farms, thus its coefficient is expected to be positive.

Current prices may not reflect expected future profitability of farming. Land values are, therefore, included as an explanatory variable since land values are determined by expected future returns from farming.[5] Increases in the value of farmland should be associated with greater anticipated returns in farming and the greater net entry of farms. Land is also an important asset in the farm portfolio, so changes in its value can affect the financial viability of farm operations. Increases in land values increase the attractiveness of farmland as an investment and should lead to greater net entry, while decreases in land values result in financial stress and lead to less net entry. The measure of land values used here is the USDA average value per acre of farm real estate deflated using the GNP deflator.

Real interest rates influence the price of credit, important to entering farmers, and can affect the degree of financial stress. Higher real interest rates may reduce net entry by making farm borrowing more expensive and increasing financial stress. The real interest rate was computed by subtracting actual inflation, as measured by the GNP deflator, from the prime rate charged by banks.

An increase in the ratio of nonfarm wages to farm income is expected to reduce net entry (increase net exit), since this will enhance the attractiveness of nonfarm employment over farm employment. Annual nonfarm income was obtained by multiplying private agricultural average gross weekly earnings in 1977 dollars by 50. Farm income was obtained by dividing net farm income in 1977 dollars by the number of farms.[6]

Estimation and Results

The results were obtained using the Yule-Walker procedure for correcting first-order autocorrelation of the error terms to achieve efficient estimates (*3,8*). An instrumental variable is created for the land value variable (because it is not exogenous and likely to be correlated with interest rate) by using predicted values from the following regression:

$$\begin{aligned} \text{Land value}_t = {} & 172.1 + 0.014\, X_t + 13.38\, R_t \\ & (28.83) \quad (0.003) \quad (5.86) \\ & + 23.99\, I_t,\ R^2{=}0.90, \\ & (6.67) \end{aligned}$$

(standard errors in parentheses) where X_t is exports of farm products deflated by the prices received by farmers index, R_t is the real interest rate, and I_t is actual inflation as measured by the GNP deflator

Table 1 presents the results of two regressions estimated for 1960-88. An F-test rejects the hypothesis that the coefficients on the economic variables are

[5]Barkley (*1*) included land values in a model that explained migration of labor out of agriculture and found a positive association.

[6]This measure does not take into account the fact that many U.S. farmers supplement farm income with income from off-farm sources.

Table 1—Regression results: annual change in number of U.S. farms, 1960-88[1]

Variable[2]	(1) Full model	(2) Reduced model
Intercept	133,486 * (28,760)	162,128 * (38,882)
Number of farms {t-1}	-.07 * (.01)	-.08 * (.01)
Ratio of prices received to prices paid by farmers {t}	53,683 * (27,282) [.93]	
Average value/acre of farm real estate (1977 dollars) {t-1}	112.81 * (28.95) [1.07]	
Real interest rate {t-1}	-3,432 * (1,395) [.18]	
Ratio of nonfarm earnings to real farm income {t-1}	2,525 (5,282)	
Durbin-Watson statistic	1.05	.47
Adjusted R^2	.90	.57

[1]These estimates were obtained by using a procedure to correct for autocorrelation of the error terms.
[2]The dependent variable is the change in the number of U.S. farms between time t and time t-1. { } = Time subscript. ( ) = Standard errors. [ ] = Elasticities computed at values. * = Significance at the 5-percent level with a one-tailed test.

jointly equal to zero (equation 4), indicating that these variables do add explanatory power to the model. Three of the four economic variables are significant and have the expected signs, with the ratio of nonfarm wages to farm income being the only nonsignificant variable. The model that includes the economic variables explains 90 percent of the variation in the change in farm numbers, while the model using only the lagged number of farms explains only 57 percent of the variation.

The estimate of α is minus 0.07 in the full model (column 1), suggesting that the number of farms falls each year by 7 percent of the difference between the number of farms and the value of F_n. The value of F_n implied is 1,906,943. This agrees with the data of figure 1, which suggested that the number of farms has been approaching 2 million asymptotically over time.

The ratio of output to input prices has a significant positive coefficient, as does the coefficient on land values. These results suggest that as the output-input price ratio and land value rises (falls), thus increasing (decreasing) the profitability of farming and its attractiveness as an investment, net entry increases (decreases). The real interest rate has a negative effect, suggesting that as real interest rates rise, the price of credit rises, deterring new entry, and possibly resulting in greater financial stress among current farmers, leading to increased exits. The final result is less net entry.

The computed elasticities are approximately equal to one for the coefficients on prices and land values, and 0.18 for the interest rate, indicating that the magnitude of the price and land value effects is much greater than that of the interest rate. An increase in the ratio of prices received by 0.10 will increase net entry by 5,368, an increase in land value of $100 will increase net entry by 11,281. An increase in the interest rate of 1 percentage point will reduce net entry by 3,432.

The data were divided into two subperiods to test for differences over time in the importance of the economic variables and in the trend found in the data. The 1960-74 period appeared to be dominated by the longrun trend toward fewer farms at a declining rate (fig. 2). During this early period, changes in farm numbers showed much less year-to-year variation than during 1975-88, perhaps suggesting that shortrun economic conditions exerted less influence than in recent years, or that economic conditions were less volatile. I estimated separate regressions over the two time periods to detect differences in the parameters between the periods.

The results of the regressions show significant differences between the two periods (table 2). A Chow test (*4*, p. 87) rejects the hypothesis that the parameters of the models, including the economic variables, are equal between the two subperiods with an F-statistic of 4.0 compared with a critical $F(6,16)$ at the 5-percent level of 2.74.

The economic variables have much greater importance in the later period, with all except the income ratio being significant. The land value coefficient is significant in the early period, but an F-test fails to reject the hypothesis that the coefficients on the economic variables are jointly equal to zero. The regression using only the lagged value of farms for 1960-88 explains 87 percent of the variation in the change in farm numbers. This reduced model has no explanatory power at all during 1975-88, however, while the model including the economic variables had an R^2 of 0.95.

Table 2—Regression by subperiod: annual change in number of U.S. farms[1]

Variable[2]	1960-74		1975-88	
	(1) Full model	(2) Reduced model	(3) Full model	(4) Reduced model
Intercept	106,897 * (37,778)	182,746 * (29,805)	702,403 * (97,807)	-37,631 (154,981)
Number of farms {t-1}	-.063 * (.012)	-.087 * (.009)	-.305 * (.041)	.004 (.065)
Prices received/ prices paid by farmers {t}	46,077 (34,469)		189,063 * (56,060) [6.29]	
Average value/ acre farmland (1977 dollars) {t}	234.03 * (99.36) [1.14]		267.04 * (38.17) [7.06]	
Real interest rate {t-1}	-2,258 (2,438)		-5,725 * (2,141) [.91]	
Ratio of nonfarm to farm income {t-1}	28,500 (23,894)		7,815 (5,303)	
Durbin-Watson statistic	1.23	.46	2.93	.90
Adjusted R^2	.98	.87	.95	.0

[1]These estimates were obtained by using a procedure to correct for autocorrelation of error terms. Columns 1 and 2 were estimated using data for 1960 to 1974, columns 3 and 4 were estimated using data for 1975 to 1988.

[2]The dependent variable is the change in the number of U.S. farms between period t and t-1. { } = Time subscripts.. ( ) = Standard errors. * = Significance at the 5-percent level using a one-tailed test.

When the economic variables are included in the 1975-88 model (column 3) the intercept and the coefficient on the lagged number of farms are significant, the value of α being 0.305 and the value of F_n about 2.3 million, while both of these parameters are nonsignificant in the reduced model (column 4). The decline in farm numbers at a decreasing rate can be detected in the 1975-88 data when one accounts for the effect of the economic variables, although the value of α is larger than that for 1960-74 (0.06 to 0.09).

The effects of prices, land values, and interest rates appear to be of much greater magnitude during 1975-88 than during 1960-74. The coefficient on land values is not statistically different between the two periods, but the proportional effect is much larger in the later period (an elasticity of 7 versus a value of 1 for the early period) because the number of farms is much smaller in the later period. The price variable had an elasticity of 6.29 during 1975-88 and was nonsignificant during 1960-74. The interest rate was also nonsignificant in the early period, with an elasticity of 0.93 in the later period. The point estimates, however, are not statistically different between the two periods. As in the 1960-88 model, the proportional effects of prices and land values during 1975-88 are about equal to each other and much larger than the effect of interest rates.

The implication of these results is that economic variables have taken on an increasingly important role in recent years in influencing changes in farm numbers while in previous years changes were mainly due to longrun structural forces. Table 3 shows coefficients of variation for the variables under study for 1960-74 and 1975-88. The large coefficients of

Table 3—Descriptive statistics by subperiod

Variable	1960-74 Mean	1960-74 Coefficient of variation	1975-88 Mean	1975-88 Coefficient of variation
Number of U.S. farms	3,054,361	13.99	2,366,450	4.99
Annual change in number of U.S. farms	-87,271	-44.57	-27,961	-68.62
Percentage change in number of U.S. farms	-2.76	-34.68	-1.19	-70.45
Ratio of prices received to prices paid by farmers	1.17	6.85	.93	13.57
Average value/acre of farm of real estate (million 1977 dollars)	439	14.97	740	22.18
Real interest rate	2.11	45.15	4.67	76.46

variation in 1975-88 indicate that prices, land values, and interest rates have shown much greater volatility during the more recent period, as have changes in farm numbers. The results suggest that the greater fluctuations in farm numbers experienced in recent years result both from greater sensitivity to economic conditions and from greater variability in economic conditions.

Implications for Modeling Farm Structure

This increasing importance of fluctuations in economic conditions has important implications for the modeling of farm structure, which is normally studied with the use of census data from a limited number of years. Estimating Markov chain models meant using this type of data which were, until recently, mainly extrapolations of historical trends. The results of this study, however, suggest that variation in economic conditions may cause the probabilities of entry and exit to vary from year to year, so the transition probabilities may not be appropriate for forecasting future farm structure when economic conditions are changing.

Some recent work has emphasized that the transition probabilities[7] estimated in Markov chain analysis of farm structure are nonstationary over time (*7, 10*). The data available have permitted only cross-sectional analyses which are not able to estimate the effects of variables such as prices, land values, and interest rates on the transition probabilities. The greater sensitivity of farm numbers to these shortrun influences makes this problem particularly acute.

Smith (*10*), for example, projects farm numbers to 1986 using Markov models estimated from 1974 to 1978 census data. His model overpredicts farm numbers from 1978 to 1986 because 1974-78 was a period when the economic environment of farming was relatively favorable (hence the probability of exit, say, was unusually low), and the model could not take into account the declines in prices and land values and increases in real interest rates that occurred in the 1980's and triggered drastic reductions in farm numbers.[8]

Until models can incorporate economic variables that vary from year to year into nonstationary analysis, results obtained from them may be of limited temporal generality, and the projections made from such models should be used and interpreted carefully. It may become possible to incorporate these variables into

[7]The probability that a farm in class i at time t will move to class j at time t+1. The classes are usually discrete size classes plus a nonfarm class which allows for entry and exit.

[8]Admittedly, a time series model such as the one estimated in this article could not have made the prediction unless it could have anticipated the changes in prices.

nonstationary Markov models as the length of the census longitudinal file (*2,7,9,10*) is increased to include additional years, allowing more variation in prices and policies.

Conclusion

This study showed by means of regression analysis that economic variables, including prices, land values, and interest rates, have a significant influence on changes in the number of farms. The influence of these variables also appears to have increased in recent years, where in earlier years the number of farms evolved in a more deterministic manner, falling at a decreasing rate. When I controlled for the influence of these variables, I found the trend in the number of farms still to be present in recent years.

It is important to understand the operation of these influences given the increasing integration of the farm sector into the general economy of the Nation and into the world economy. An understanding of these influences is essential for guiding discussions about farm policy, where influences on farm numbers are often a central point. Future models of farm structure and projections of the number of farms should consider both the longrun trends and the influence of shortrun fluctuations in economic variables, since these two effects operate in tandem to determine the number of farms.

Appendix

The data on farm numbers are annual estimates produced by USDA and published in August issues of *Crop Production* (*11*) and various statistical bulletins (*12,13,14*). The estimates are linked to the census counts, and for noncensus years the number of farms is estimated using information from the USDA June enumerative survey. While the accuracy of the estimates may be open to question, the numbers do represent USDA's official estimates of the number of farms. The total number of farms does not tell the complete story about farm structure, since it does not reveal anything about relative changes in size classes.

A change in the definition of what is considered to be a farm occurred in 1974, resulting in a discontinuity in the published series. A farm had previously been defined as a place of more than 10 acres and at least $50 of sales of agricultural products or a place of 10 acres with sales of at least $250. The cutoff was changed to $1,000 regardless of acreage under the new definition. Numbers were published for 4 years under both the new and old definitions, and a comparison revealed that the new and old estimates differed by about 245,000 farms. The old and new series were spliced together by subtracting 245,000 farms from all pre-1975 numbers. While this works on the unlikely assumption that the number of farms with sales of less than $1,000 remained constant over the entire period, a more sophisticated adjustment scheme was not possible. Another distortion results from inflation which tends to push more very small farming operations into the "farm" definition as prices rise, increasing the value of sales for a given quantity of production.

References

1. Barkley, Andrew. "The Determinants of the Migration of Labor Out of Agriculture: 1940-85." Dept. of Agricultural Economics Staff Paper No. 90-1, Kansas State University, Manhattan, Oct. 1988.

2. Edwards, Clark, Matthew G. Smith, and R. Neal Peterson. "The Changing Distribution of Farms by Size: A Markov Analysis," *Journal of Agricultural Economics Research*, Vol. 37, No. 4, 1985, pp. 1-16.

3. Gallant, A.R., and J.J. Goebel. "Non-Linear Regression With Autoregressive Errors," *Journal of the American Statistical Association*, Vol. 71, 1976, pp. 961-7.

4. Kennedy, Peter. *A Guide to Econometrics*. Cambridge: MIT Press, 1987.

5. Lin, William, George Coffman, and J.B. Penn. *U.S. Farm Numbers, Sizes, and Related Structural Dimensions: Projections to the Year 2000*. TB-1625. U.S. Dept. Agr., Econ., Stat., Coop. Serv., July 1980.

6. Office of Technology Assessment, U.S. Congress. *Technology, Public Policy, and the Changing Structure of American Agriculture*. OTA-F-285. 1986.

7. Peterson, R. Neal. "A Single Equation Approach to Estimating Nonstationary Markov Matrices: The Case of U.S. Agriculture," U.S. Dept. Agr., Econ. Res. Serv., forthcoming.

8. SAS Institute, Inc. *SAS/ETS User's Guide, Version 5 Edition.* Cary, NC, 1984.

9. Smith, Matthew G. "Entry, Exit, and the Age of Distribution of Farm Operators, 1974-1982," *Journal of Agricultural Economics Research.* Vol. 39, No. 4, 1987, pp. 2-11.

10. _____. *A Conditional Approach to Projecting Farm Structure.* ERS Staff Report AGES880208. U.S. Dept. Agr., Econ. Res. Serv., May 1988.

11. U.S. Department of Agriculture, National Agricultural Statistics Service. *Crop Production.* August 1984-88.

12. _____, Statistical Reporting Service. *Farms and Land in Farms: Final Estimate by States, 1969-75.* SB-594. Dec. 1977.

13. _____. *Farms: Revised Estimates 1959-70.* SB-507. Jan. 1973.

14. _____. *Number of Farms 1910-1959, Land in Farms 1950-1959 by States.* SB-316. June 1962.

The Deterministic Equivalents of Chance-Constrained Programming

C.S. Kim, Glenn Schaible, and Eduardo Segarra

Abstract. *Three concepts combine to show both the feasibility and desirability of incorporating probability within programming models. First, the reliability of estimates obtained by using Chebyshev's inequality increases as variation, measured by the coefficient of variation, declines. Second, the coefficient of variation can be substantially reduced by the use of the mean and variance of a truncated normal distribution. Third, chance-constrained programming can be converted into deterministic equivalent quadratic programming by using the parameters of a truncated normal distribution.*

Keywords. *Chance-constrained programming, quadratic risk programming, truncated normal distribution, Chebyshev's inequality.*

The developments of the past three decades in the theory of choice under risk or uncertainty have followed the expected mean-variance approach. The decisionmaker is assumed to select among alternative activities on the basis of a utility function defined in terms of the expected mean and variance of the portfolio return. Accordingly, much has been written about quadratic risk programming (*5, 7, 13, 16*).[1] Researchers have not paid enough attention, however, to the burdensome data requirements and mathematical complexity associated with the risk and uncertainty pertaining to input-output coefficients *(9)*.

An activity analysis model usually optimizes some objective function subject to linear constraints. Coefficients for both the objective function and the constraints are assumed to be known with certainty. Chance-constraint programming (CCP), originally proposed by Charnes and Cooper (*4*), makes use of individual probabilistic constraints. A probability is attached to the linear constraint in such models. The probabilistic constraint is subject to some predetermined critical level, and its coefficients are assumed to be randomly distributed. Probabilistic constraints of this type often appear in decision analysis in the form of a safety-first rule in portfolio selection problems (*6,11,12,14,15*), or in problems associated with feed-grain mixture (*5*).

Many researchers, therefore, have attempted to convert probabilistic constraints into deterministic equivalents under various assumptions. Examples include the works of Charnes (*3*), Pyle and Turnovsky (*11*), Roy (*12*), Telser (*15*), Paris and Easter (*9*), Sengupta (*14*), and Atwood (*1*). Paris and Easter, and Pyle and Turnovsky obtain their criterion based on the normality assumption with respect to the random variable. Roy and Telser, in contrast, obtain deterministic equivalents of a probabilistic constraint by using Chebyshev's inequality, which does not require any knowledge about the probabilistic density function of a random variable. Sengupta claims that estimates obtained from the use of Chebyshev's inequality may "sometimes" be very inefficient, in the sense that they may provide very rough approximations for the actual probability when the distribution of the random variable is known. However, Sengupta failed to indicate when the use of Chebyshev's inequality provides inefficient estimates of the actual probability.

Recently, Atwood and others (*1,2*) rejected the use of Chebyshev's inequality on the grounds that the estimates are too conservative. They proposed the use of lower partial moments (LPM) to obtain a deterministic equivalent of a probabilistic condition. The LPM approach uses the concept of a truncated distribution, which Sengupta suggested to improve reliability of estimates. However, the LPM approach makes use of the parameters of a complete normal distribution. Unknown is how much the reliability of estimates is improved by use of a truncated distribution using parameters of a complete normal distribution over the use of the Chebyshev inequality.[2] In this paper, we first demonstrate that the reliability of estimates obtained by making use of Chebyshev's inequality increases as the coefficient of variation decreases. Second, we demonstrate that, under certain conditions, use of a truncated normal distribution does not necessarily improve the reliability of estimates, and that estimates obtained from the use of Chebyshev's in-

Kim and Schaible are economists with the Resources and Technology Division, ERS; Eduardo Segarra is an assistant professor with the Department of Agricultural Economics, Texas Tech University, Lubbock. The authors thank Kuo S. Huang for his review of early drafts of this manuscript.

[1]Italicized numbers in parentheses cite sources listed in the References section at the end of this article.

[2]The reliability of an estimate is defined as the probability of the absolute value of the difference between the estimate and its true value as being less than a predetermined small number.

equality are generally reliable if the analysis includes the use of the mean and variance of a truncated normal distribution. Third, we show how one can convert the CCP problem into a deterministic equivalent quadratic programming (DEQP) problem using Chebyshev's inequality. Finally, we discuss the properties of DEQP solutions.

A Complete Normal Distribution and Chebyshev's Inequality

The normality assumption has been widely used in economic literature, and its use is justified for many cases due to the Central Limit Theorem. However, without any knowledge about the probability density function of a random variable x, we can consider Chebyshev's inequality for any $h > 0$, such that:

$$\Pr[|x-\mu| \leq h\sigma] \geq 1 - \frac{1}{h^2}, \tag{1}$$

where μ and σ are the mean and standard deviation of the complete normal distribution.

For any symmetric distribution, the quantity $Q^*(h)$, which is the probability assigned to the interval where the random variable x is defined as $\{x: (\mu - h\sigma) < x \leq \infty\}$, may be represented as:

$$\begin{aligned} Q^*(h) &\geq 0.5 + 0.5\left[1 - \frac{1}{h^2}\right] \\ &\geq 1 - 1/(2h^2). \end{aligned} \tag{2}$$

It is clear from equations 1 and 2 that the probability of an observed value for a random variable x in the interval $(x: (\mu - h\sigma) < x \leq \infty)$ approaches one as h increases infinitely, and that the probability increases faster for symmetric distributions than for non-symmetric distributions.

Table 1 shows probabilities estimated with equation 2 for h = 1, 2, . . . , 5. By making use of Chebyshev's inequality, we estimated that probabilities are relatively reliable estimates for a random variable with a small coefficient of variation (σ/μ), and that the reliability increases as h increases.[3] Table 1 indicates that for $h > 4$, the probabilistic constraint of a chance-constrained programming problem can be converted into a deterministic equivalent by making use of Chebyshev's inequality.

[3]Parzen (*10*) showed for a two-sided confidence level, $\Pr(|X-h\mu|) \geq \beta$, for a complete normal distribution, that estimates obtained using Chebyshev's inequality converge to actual probability as h increases.

Table 1—Estimated probabilities $\Pr(x > \mu\text{-}h\sigma)$ for a complete normal distribution and the use of Chebyshev's inequality

h	$\phi(h)^1$	$1 - \frac{1}{2h^2}$	Reliability of estimates with Chebyshev inequality
			Percent
1	0.8413	0.5000	59.43
2	.9772	.8750	89.54
3	.9987	.9444	94.56
4	1.0000	.9688	96.88
5	1.0000	.9800	98.00

[1]ϕ is the CDF of N (0,1).

A Truncated Normal Distribution and Chebyshev's Inequality

There exist cases in which economic variables are not defined over the entire range of a normal distribution. For instance, water applied in the production of irrigated cotton or the final mix of feed grains in livestock production cannot be negative, indicating that the random variable, say x, would have a truncated normal distribution to the left at x = 0.

When we assume that x has a truncated normal distribution to the left at τ, the probability density function of x is then defined as follows (*8*):

$$f(x) = \begin{cases} 0 & \text{if } x < \tau \\ \dfrac{K}{\sigma\sqrt{2\pi}} \exp\left[-\dfrac{1}{2}\left(\dfrac{x-\mu}{\sigma}\right)^2\right] & \text{if } x \geq \tau, \end{cases} \tag{3}$$

where

$$K = \frac{1}{1 - \phi(\frac{\tau-\mu}{\sigma})},$$

ϕ is the cumulative density function (CDF) of the normal distribution with mean zero and unity variance, N(0,1), and μ and σ are the mean and standard deviation of a complete normal distribution of x.[4]

To compare estimates from a truncated normal distribution and estimates obtained by using Chebyshev's inequality, we define the quantity Q(h), for any $h > 0$,

[4]Meyer (*8*) erroneously defined K as the quantity $1 - \phi(\frac{\tau-\mu}{\sigma})$, rather than the quantity $K = 1/[1 - \phi(\frac{\tau-\mu}{\sigma})]$.

as the probability assigned to the interval where the random variable x is $\{x: (\mu - h\sigma) < x \leq \infty\}$. For the truncated normal probability density function (with mean μ and standard deviation σ of the complete normal distribution), the quantity Q(h) is given by:

$$Q(h) = \frac{K}{\sigma\sqrt{2\pi}} \int_{(\mu-h\sigma)}^{+\infty} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx. \tag{4}$$

Equation 4 can be compactly rewritten as:

$$Q(h) = \frac{1 - \phi(-h)}{1 - \phi(\frac{\tau-\mu}{\sigma})} = \frac{\phi(h)}{\phi(\frac{\mu-\tau}{\sigma})}. \tag{5}$$

It is easy to see that Q(h) in equation 5 becomes 1 for $\tau = (\mu - h\sigma)$. In cases where $(\mu-\tau)/\sigma \geq 3.9$ so that $\phi[(\mu-\tau)/\sigma] = 1$, equation 5 becomes $Q(h) = \phi(h)$, which is the probability estimated from a complete normal distribution. That is, use of a truncated normal distribution does not improve the reliability of estimates if $(\mu-\tau)/\sigma$ is greater than or equal to 3.9.

We will show (equation 22 and footnote 5) that $(\mu-\tau)/\sigma \geq (1-\beta)^{-1/2}$, where β is the required minimum probability of success. For $\beta = 0.95$, a traditional criterion in decision analysis, $(1-\beta)^{-1/2}$ equals 4.4721 which is greater than 3.9. Use of a truncated normal distribution, therefore, does not necessarily improve reliability of estimates.

So far, the mean and standard deviation of a complete normal distribution are used for the measurement of the coefficient of variation in equation 5. However, the expected value E(x) and variance V(x) of a random variable x, which has a truncated normal distribution as given in equation 3, are expressed as follows (see appendix):

$$E(x) = \mu + D, \tag{6}$$

and

$$V(x) = \sigma^2 + D[(\tau - \mu) - D], \tag{7}$$

$$\text{where } D = \frac{\sigma}{\sqrt{2\pi}\left[1 - \phi\left(\frac{\tau-\mu}{\sigma}\right)\right]} \exp\left[-\tfrac{1}{2}\left(\frac{\tau-\mu}{\sigma}\right)^2\right].$$

In equations 6 and 7, $D > 0$, and therefore, $E(X) > \mu$, $V(X) < \sigma^2$, and $\sigma/\mu > [V(X)]^{1/2}/E(X)$. That is, the coefficient of variation is reduced when the mean and standardized deviation of a truncated normal distribution are substituted for those of a complete normal distribution, respectively. However, we have shown that the reliability of estimates obtained by making use of Chebyshev's inequality increases as the coefficient of variation is reduced. Therefore, parameter estimates of a chance-constrained problem can be improved by making use of the mean and variance of the truncated normal distribution.

Deterministic Equivalent Quadratic Programming

A probabilistic constraint can be converted into a deterministic equivalent using Chebyshev's inequality. Consider a chance-constrained programming problem such as the one used by Chen.

$$\text{Minimize} \quad C'X, \tag{8}$$

$$\text{subject to:} \quad \Pr(P'X \geq d) \geq \beta \tag{9}$$

$$AX \geq b \tag{10}$$

$$X \geq 0, \tag{11}$$

where C is an (n x 1) vector of cost coefficients; X is an (n x 1) vector of choice variables; P is an (n x 1) vector of stochastic variables, which has a truncated distribution to the left at zero, with a truncated normal mean vector $\overline{P}$ and variance-covariance matrix W; d is a prespecified constant, β is the required minimum probability of success; A is an (m x n) technical coefficient matrix; and b is an (m x 1) vector of minimum resource requirements.

To formulate a deterministic equivalent constraint, the probabilistic constraint 9 can be rewritten as:

$$\begin{aligned} \Pr(P'X \geq d) &= 1 - \Pr(P'X \leq d) \\ &= 1 - \Pr[(\overline{P}'X - P'X) \geq (\overline{P}'X - d)] \\ &\geq 1 - \Pr[(|P'X - \overline{P}'X|) \geq (\overline{P}'X - d)] \\ &\geq 1 - \frac{X'WX}{(\overline{P}'X - d)^2} \text{ Chebyshev's inequality,} \\ &\geq \beta \end{aligned} \tag{12}$$

or equivalently,

$$\Pr(P'X \geq d) \geq 1 - \frac{X'WX}{(\overline{P}'X - d)^2} \geq \beta. \tag{13}$$

The coefficient of variation clearly must be smaller for a large required probability of success.

Minimum costs that satisfy the probabilistic constraint 9 are attained at $\Pr(P'X \geq d) = \beta$, or more stringently at:

$$1 - \frac{X'WX}{(\overline{P}'X - d)^2} = \beta,$$

in equation 13. Consequently, the problem of finding minimum costs, $C'X = K$, subject to the probabilistic constraint 9, is equivalent to the problem of finding the minimum of:

$$(1 - \beta) - \frac{X'WX}{(\overline{P}'X - d)^2} \geq 0, \tag{14}$$

at any given cost K. This can be verified by comparing the Kuhn-Tucker conditions from each optimization problem. The minimization of equation 14 is equivalent to the maximization of:

$$\overline{P}'X + \frac{1}{2d}(X'MX) \leq \frac{d}{2}, \tag{15}$$

where $M = [W(1 - \beta)^{-1} - \overline{P}\,\overline{P}']$ is a negative semidefinite matrix.

When the equality holds in equation 15 for any level of required success, β, the probability constraint 9 is met at minimum cost, say K_β. In cases where the inequality holds, the probability constraint 9 is satisfied, but at a higher cost than K_β. Consequently, the following deterministic equivalent QP (DEQP) problem can be formulated, which is equivalent to the CCP problem in equations 8 through 11.

$$\text{Maximize} \quad \overline{P}'X + \frac{1}{2}(X'MX) \leq \frac{d}{2} \tag{16}$$

$$\text{subject to:} \quad C'X = K \qquad 0 < K \leq K_\beta \tag{17}$$

$$AX \geq b \tag{18}$$

$$X \geq 0. \tag{19}$$

The optimal choice variables X satisfying the CCP problem in equations 8 through 11 may be approximated by solving the DEQP problem in equations 16 through 19 by simply increasing cost (K) parametrically until the objective value approaches d/2.

Properties of DEQP Solutions

To derive the properties of DEQP solutions, we rewrote the objective function in equation 16 as:

$$\begin{aligned} Z &= \overline{P}'X + (1/2d)X'[(1 - \beta)^{-1}W - \overline{P}\,\overline{P}']X \\ &= \overline{P}'X + \frac{(1-\beta)^{-1}}{2d} X'WX - (1/2d)\, X'\overline{P}\,\overline{P}'X \\ &\leq \frac{d}{2}. \end{aligned} \tag{20}$$

Multiplying both sides of the inequality in equation 20 by 2d, we can obtain the following:

$$2d\,\overline{P}'X + (1-\beta)^{-1}X'WX - X'\overline{P}\,\overline{P}'X \leq d^2. \tag{21}$$

From equation 21, we can obtain the following inequality:

$$\frac{\overline{P}'X - d}{(X'WX)^{1/2}} \geq (1-\beta)^{-1/2}, \tag{22}$$

or equivalently:

$$\overline{P}'X - (1-\beta)^{-1/2}(X'WX)^{1/2} \geq d. \tag{22'}$$

When the condition given in equation 22′ is met, it can be said that one is β-percent confident, and that $P'X$ will be greater than or equal to a predetermined constant, d.[5]

To further investigate the properties of DEQP solutions, consider the Lagrangian equation associated with the DEQP problem in equations 16 through 19 such that:

$$\begin{aligned} F(X, \lambda_1, \lambda_2) = \overline{P}'X + (1/2d)X'\,[(1-\beta)^{-1}W - \overline{P}\,\overline{P}']X \\ - \lambda_1(K - C'X) - \lambda'_2(b - AX), \end{aligned} \tag{23}$$

where λ_1 is a Lagrangian multiplier, such that $\lambda_1 \gtrless 0$, and λ_2 is an (m x 1) vector of positive Lagrangian multipliers.

Part of the Kuhn-Tucker conditions are expressed as:

$$\begin{aligned} \frac{\partial F}{\partial X} = \overline{P} + \frac{(1-\beta)^{-1}}{d} WX - (1/d)\overline{P}\,\overline{P}'X + \\ \lambda_1 C + \lambda'_2 A \leq 0 \end{aligned} \tag{24}$$

[5]The condition in equation 22′ is more stringent than the one imposed by negative semidefiniteness. Because $M = [(1-\beta)^{-1}W - \overline{P}\,\overline{P}']$ is a negative semidefinite matrix, $X'MX = (1-\beta)^{-1}X'WX - X'\overline{P}\,\overline{P}'X \leq 0$, and therefore:

$$\frac{X'\overline{P}\,\overline{P}'X}{X'WX} \geq (1-\beta)^{-1},$$

or equivalently, $\overline{P}'X - (1-\beta)^{-1/2}(X'WX)^{1/2} \geq 0$, which is less stringent than the condition expressed in the inequality 22′.

$$\left(\frac{\partial F}{\partial X}\right)X = \overline{P}'X + \frac{(1-\beta)^{-1}}{d} X'WX - (1/d)X'\overline{P}\,\overline{P}'X$$
$$+ \lambda_1 C'X + \lambda'_2 AX = 0 \quad (25)$$

$$X \geq 0. \quad (26)$$

There exists a dual programming problem associated with the primal programming problem given in equations 16 through 19. The objective function to be minimized in a dual programming problem is obtained by simply subtracting equation 25 from the Lagrangian equation 23, represented by:

$$G(X, \lambda_1, \lambda_2) = -\lambda_1 K - \lambda'_2 b$$
$$+ (1/2d)X'\overline{P}\,\overline{P}'X - \frac{(1-\beta)^{-1}}{2d} X'WX. \quad (27)$$

The objective value of the primal programming problem, equation 16, equals the objective value of the dual programming problem, equation 27, at optimum. Equating equations 16 and 27 results in the following:

$$\overline{P}'X + \frac{(1-\beta)^{-1}}{d} X'WX - (1/d)X'\overline{P}\,\overline{P}'X$$
$$= -\lambda_1 K - \lambda'_2 b. \quad (28)$$

By multiplying both sides of the equality in equation 28 by d, and with minor manipulation, we obtain:

$$d[\overline{P}'X - \lambda_1 K - \lambda'_2 b] = 2d\overline{P}'X + (1-\beta)^{-1}X'WX$$
$$- X'\overline{P}\,\overline{P}'X \quad (29)$$
$$\leq d^2 \quad \text{(from equation 21)},$$

or equivalently,

$$\overline{P}'X \leq d + \lambda_1 K + \lambda'_2 b. \quad (30)$$

Equation 30 shows that the mean value of the random variable P'X must be less than or equal to the predetermined value d plus the sum of the opportunity costs of expenditures and the opportunity costs of other resources.

Combining conditions in equations 22′ and 30 reveals that:

$$d + (1-\beta)^{-1/2}(X'WX)^{1/2} \leq \overline{P}'X \leq d + \lambda_1 K + \lambda'_2 b, \quad (31)$$

or equivalently,

$$\overline{P}'X - (1-\beta)^{-1/2}(X'WX)^{1/2} \geq d \geq \overline{P}'X - \lambda_1 K - \lambda'_2 b. \quad (32)$$

Both sides of the first inequality sign in equation 32 are identical with the condition given in equation 22′. However, the condition imposed by equation 32 is more restrictive than the one given in equation 22′ by requiring an additional condition given in equation 30.

Conclusions

Chebyshev's inequality often has been used to convert a probabilistic constraint in a chance-constrained programming problem into a deterministic constraint. Researchers, however, have criticized the use of Chebyshev's inequality which sometimes provides very rough approximations for the actual probability. We have shown that the use of Chebyshev's inequality provides relatively very good approximations for the actual probability when the coefficient of variation is relatively very small. We also have shown that the use of the mean and variance of a truncated normal distribution reduces the size of the coefficient of variation, compared with the case of using the mean and variation of a complete normal distribution. We have also demonstrated how a chance-constrained programming problem can be converted into a deterministic equivalent quadratic programming problem.

Appendix 1— Mean and Variance of a Truncated Normal Distribution

$$E(x) = \int_{T}^{+\infty} xf(x)dx = \int_{T}^{+\infty} \frac{X\left[\dfrac{1}{1 - \phi(\frac{T-\mu}{\sigma})}\right]}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{X-\mu}{\sigma})^2} dx .$$

With $S = (\frac{X-\mu}{\sigma})$, and therefore $\sigma dS = dx$, then:

$$E(x) = \frac{1}{1 - \phi(\frac{T-\mu}{\sigma})} \int_{\frac{T-\mu}{\sigma}}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} (\mu + \sigma S) e^{-S^2\, 2} \sigma dS$$

$$= \frac{1}{1 - \phi(\frac{T-\mu}{\sigma})} \left[\mu \int_{\frac{T-\mu}{\sigma}}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-S^2\, 2} dS + \sigma \int_{\frac{T-\mu}{\sigma}}^{+\infty} \frac{S}{\sqrt{2\pi}} e^{-S^2\, 2} dS \right]$$

$$= \frac{1}{1 - \phi(\frac{T-\mu}{\sigma})} \left[\mu \left[1 - \phi(\tfrac{T-\mu}{\sigma}) \right] + \frac{\sigma}{\sqrt{2\pi}} \left(-e^{-S^2\, 2} \Big|_{\frac{T-\mu}{\sigma}}^{+\infty} \right) \right]$$

$$= \mu + \left[\frac{1}{1 - \phi(\frac{T-\mu}{\sigma})} \right] \frac{\sigma}{\sqrt{2\pi}} \left(e^{-\frac{1}{2}(\frac{T-\mu}{\sigma})^2} \right)$$

$$E(x^2) = \int_{T}^{+\infty} x^2 f(x)dx = \int_{T}^{+\infty} \frac{X^2 \left[\dfrac{1}{1 - \phi(\frac{T-\mu}{\sigma})}\right]}{\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{X-\mu}{\sigma})^2} dx$$

$$= \frac{1}{\sqrt{2\pi}}\left(\frac{1}{1-\phi(\frac{T-\mu}{\sigma})}\right)\int_{\frac{T-\mu}{\sigma}}^{+\infty} (\mu + \sigma S)^2 e^{-(\frac{1}{2})S^2} dS$$

$$= \frac{1}{1-\phi(\frac{T-\mu}{\sigma})}\left[\int_{\frac{T-\mu}{\sigma}}^{+\infty} \frac{\mu 2}{\sqrt{2\pi}} e^{-(\frac{1}{2})S^2} dS + 2\mu\sigma \int_{\frac{T-\mu}{\sigma}}^{+\infty} \frac{S}{\sqrt{2\pi}} e^{-(\frac{1}{2})S^2} dS\right.$$

$$\left. + \sigma^2 \int_{\frac{T-\mu}{\sigma}}^{+\infty} \frac{S^2}{\sqrt{2\pi}} e^{-(\frac{1}{2})S^2} dS\right]$$

$$= \frac{1}{1-\phi(\frac{T-\mu}{\sigma})}\left[\mu^2[1-\phi(\frac{T-\mu}{\sigma})]\right] + \frac{1}{1-\phi(\frac{T-\mu}{\sigma})} \bullet \frac{2\mu\sigma}{\sqrt{2\pi}}\left(-e^{-S^2/2} dS \Big|_{\frac{T-\mu}{\sigma}}^{+\infty}\right)$$

$$+ \frac{1}{1-\phi(\frac{T-\mu}{\sigma})}\left[\frac{\sigma^2}{\sqrt{2\pi}} \int_{\frac{T-\mu}{\sigma}}^{+\infty} S^2 e^{-(\frac{1}{2})S^2} dS\right]$$

$$= \mu^2 + \frac{1}{1-\phi(\frac{T-\mu}{\sigma})} \bullet \frac{2\mu\sigma}{\sqrt{2\pi}}\left(e^{-\frac{1}{2}(\frac{T-\mu}{\sigma})^2}\right) + \frac{1}{1-\phi(\frac{T-\mu}{\sigma})}\left[\frac{\sigma^2}{\sqrt{2\pi}} \int_{\frac{T-\mu}{\sigma}}^{+\infty} S^2 e^{-(\frac{1}{2})S^2} dS\right]$$

Let $v = -e^{-S^2/2}$ $\qquad$ $dv = Se^{-S^2/2}dS$

$u = S$ $\qquad$ $du = dS.$

Then the last term of the above equation can be rewritten as follows:

$$\frac{1}{1-\phi(\frac{T-\mu}{\sigma})}\left[\frac{\sigma^2}{2\pi} \int_{\frac{T-\mu}{\sigma}}^{+\infty} S^2 e^{-(\frac{1}{2})S^2} dS\right] = \frac{1}{1-\phi(\frac{T-\mu}{\sigma})}\left[\frac{\sigma^2}{\sqrt{2\pi}}\left(-Se^{-S^2/2}\Big|_{\frac{T-\mu}{\sigma}}^{+\infty}\right) \frac{\sigma}{\sqrt{2\pi}} \int_{\frac{T-\mu}{\sigma}}^{+\infty} e^{-(\frac{1}{2})S^2} dS\right]$$

$$= \frac{1}{1-\phi(\frac{T-\mu}{\sigma})} \bullet \frac{\sigma^2}{\sqrt{2\pi}} \bullet (\frac{T-\mu}{\sigma}) \bullet \left(e^{-\frac{1}{2}(\frac{T-\mu}{\sigma})^2} \right)$$

$$+ \frac{1}{1-\phi(\frac{T-\mu}{\sigma})} \left[\sigma^2 \int_{\frac{T-\mu}{\sigma}}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-s^2/2} ds \right]$$

$$= \frac{1}{1-\phi(\frac{T-\mu}{\sigma})} \bullet \frac{\sigma^2}{\sqrt{2\pi}} \bullet (\frac{T-\mu}{\sigma}) \bullet \left(e^{-\frac{1}{2}(\frac{T-\mu}{\sigma})^2} \right) + \frac{1}{1-\phi(\frac{T-\mu}{\sigma})} \bullet \sigma^2 \left[1-\phi(\frac{T-\mu}{\sigma}) \right]$$

Therefore:

$$E(x^2) = \mu^2 + \sigma^2 + \frac{1}{1-\phi(\frac{T-\mu}{\sigma})} \bullet \frac{2\mu\sigma}{\sqrt{2\pi}} \left(e^{-\frac{1}{2}(\frac{T-\mu}{\sigma})^2} \right) + \frac{1}{1-\phi(\frac{T-\mu}{\sigma})} \bullet \frac{\sigma^2}{\sqrt{2\pi}} \bullet (\frac{T-\mu}{\sigma}) \bullet \left(e^{-\frac{1}{2}(\frac{T-\mu}{\sigma})^2} \right)$$

Consequently, we have the following:

$$V(x) = E(x^2) - [E(x)]^2$$

$$= \mu^2 + \sigma^2 + \frac{1}{1-\phi(\frac{T-\mu}{\sigma})} \bullet \left(e^{-\frac{1}{2}(\frac{T-\mu}{\sigma})^2} \right) \bullet \left[\frac{2\mu\sigma}{\sqrt{2\pi}} + \frac{\sigma^2}{\sqrt{2\pi}} (\frac{T-\mu}{\sigma}) \right]$$

$$- \left[\mu + \left[\frac{1}{1-\phi(\frac{T-\mu}{\sigma})} \right] \frac{\sigma}{\sqrt{2\pi}} \left(e^{-\frac{1}{2}(\frac{T-\mu}{\sigma})^2} \right) \right]^2$$

$$= \mu^2 + \sigma^2 + \frac{1}{1-\phi(\frac{T-\mu}{\sigma})} \bullet \left(e^{-\frac{1}{2}(\frac{T-\mu}{\sigma})^2} \right) \bullet \left[\frac{(\mu+T)\sigma}{\sqrt{2\pi}} \right]$$

$$- \mu^2 - \left[\frac{1}{1-\phi(\frac{T-\mu}{\sigma})} \right] \frac{2\mu\sigma}{\sqrt{2\pi}} \left(e^{-\frac{1}{2}(\frac{T-\mu}{\sigma})^2} \right) - \left[\frac{1}{1-\phi(\frac{T-\mu}{\sigma})} \bullet \frac{\sigma}{\sqrt{2\pi}} \left(e^{-\frac{1}{2}(\frac{T-\mu}{\sigma})^2} \right) \right]^2$$

Therefore:

$$V(x) = \sigma^2 + \frac{(\tau - \mu)}{1 - \phi(\frac{\tau-\mu}{\sigma})} \bullet \left[\frac{\sigma}{\sqrt{2\pi}}\right] \bullet \left(e^{-\frac{1}{2}(\frac{\tau-\mu}{\sigma})^2}\right) - \left[\frac{1}{1 - \phi(\frac{\tau-\mu}{\sigma})} \bullet \frac{\sigma}{\sqrt{2\pi}} \bullet \left(e^{-\frac{1}{2}(\frac{\tau-\mu}{\sigma})^2}\right)\right]^2.$$

References

1. Atwood, Joseph. "Demonstration of the Use of Lower Partial Moments to Improve Safety-First Probability Limits," *American Journal of Agricultural Economics.* Vol. 67, 1985, pp. 787-93.

2. Atwood, Joseph, Myles J. Watts, and Glenn A. Helmers. "Chance-Constrained Financing as a Response to Financial Risk," *American Journal of Agricultural Economics.* Vol. 70, 1988, pp. 79-89.

3. Charnes, A. "Deterministic Equivalents for Optimizing and Satisfying Under Chance-Constraints," *Operations Research.* Vol. 11, 1963, pp. 18-39.

4. Charnes, A., and W.W. Cooper. "Chance-Constrained Programming," *Management Science.* Vol. 6, 1959, pp. 73-79.

5. Chen, Joyce. "Quadratic Programming for Least-Cost Feed Formulations Under Probabilistic Protein Constraints," *American Journal of Agricultural Economics.* Vol. 55, 1973, pp. 175-83.

6. Day, R.H., D.J. Aigner, and K.R. Smith. "Safety Margins and Profit Maximization in the Theory of the Firm," *Journal of Political Economics.* Vol. 79, 1971, pp. 1,293-1,301.

7. Freund, R.J. "The Introduction of Risk into a Programming Model," *Econometrica.* Vol. 24, 1956, pp. 253-63.

8. Meyer, Paul L. *Introductory Probability and Statistical Applications.* Addison-Wesley Publishing Co., 1966.

9. Paris, Q., and C.D. Easter. "A Programming Model with Stochastic Technology and Prices: The Case of Australian Agriculture," *American Journal of Agricultural Economics.* Vol. 67, 1985, pp. 120-29.

10. Parzen, Emanuel. *Modern Probability Theory and Its Applications.* New York: John Wiley and Sons, Inc., 1960.

11. Pyle, David H., and S.J. Turnovsky. "Safety-First and Expected Utility Maximization in Mean-Standard Deviation Portfolio Analysis," *Review of Economics and Statistics.* Vol. 52, 1970, pp. 75-81.

12. Roy, A.D. "Safety-First and the Holding of Assets," *Econometrica.* Vol. 20, 1952, pp. 431-49.

13. Scott, J.T., and C.B. Baker. "A Practical Way to Select an Optimum Farm Plan Under Risk," *American Journal of Agricultural Economics.* Vol. 54, 1972, pp. 657-60.

14. Sengupta, J.K. "Safety-First Rules Under Chance-Constrained Linear Programming," *Operations Research.* Vol. 17, 1969, pp. 112-32.

15. Telser, L. "Safety-First and Hedging," *Review of Economic Studies.* Vol. 22, 1955, pp. 1-16.

16. Wiens, T.B. "Peasant Risk Aversion and Allocative Behavior: A Quadratic Programming Experiment," *American Journal of Agricultural Economics.* Vol. 58, 1976, pp. 629-35.

Combining Survey and Other Data To Estimate Agricultural Land Values

Charles H. Barnard and Danny Pfeffermann

Abstract. *Combining survey, census, and administrative data improves the precision of survey estimates of mean agricultural land values. A components-of-variance model is developed and applied to cropland value data for the Corn Belt. Performance of the model compared with other procedures is tested using cross-validation techniques. Results indicate that use of the proposed estimator would improve upon the USDA estimators at both the State and strata levels. At the strata level particularly, the improvements may be very substantial.*

Keywords. *Agricultural land values, components of variance, cross-validation, small areas, survey data.*

The Economic Research Service (ERS) estimates farm real estate value for 48 States and the United States (*24*).[1] These estimates are derived from reports obtained in the Agricultural Land Values Survey (ALVS) from sampled farmers (*14,26*). The need for improved State-level values prompted an examination of alternative data sources and alternative estimators. This paper shows how data available from sources other than the ALVS, and known at the county level, can be combined with ALVS data to improve the statistical precision of farmland value estimates. The methods described here may be useful for improving the precision of other agricultural statistics.

Small area (or small domain) estimation provides the foundation for a new estimator that combines data from the ALVS and other sources. The basis for the new estimator is a prediction model that relates the individual farmers' reports to a set of regressor variables and a set of county and State effects. The regressor variables measure known aggregate county characteristics, while the State and county effects represent specific influences not accounted for by the regressor variables. In view of the large number of counties and the small sample sizes realized in all the counties and some of the States, the county and State effects are considered random, giving rise to a mixed linear model with fixed regression coefficients and random components of variance. The use of linear models with random effects is a common practice in small domain estimation. The form of our model and the regressors included are chosen in order to best predict the small domain means of the target variable (the farmland values in the present case) and not necessarily to represent causal relationships with a substantive interpretation. Indeed, while regression analysis has been used extensively to identify causal factors explaining the value of farmland, regression techniques have not been used to yield improved estimators (predictors) of mean farmland values.

We show how the mixed linear model, in the context of small domain estimation, can potentially improve upon current USDA procedures. Data from existing sources, measuring county characteristics that are believed to affect the farmland values, are selected as regressor variables. Actual computation of the new estimator and its standard error (which we describe) permits an assessment of model performance and a comparison with USDA and other related estimators. The results of that study, conducted using cross-validation techniques, show that the new estimator, in most instances, substantially improves upon the estimators used by USDA, particularly at the strata level.

Small Domain Estimation and the Mixed Linear Model

The problem underlying the computation of the farmland value indexes may be traced to the framework of survey sampling theory. A survey population of all the farmers in the United States is divided into production regions, States, and counties. The counties are grouped into homogeneous strata and a random sample of farmers is drawn from every stratum using a probability sampling plan.[2] If the samples within the various strata were sufficiently large, one could

Barnard is an agricultural economist with the Agriculture and Rural Economy Division, ERS. Pfeffermann is senior lecturer, Dept. of Statistics, Hebrew University, Jerusalem. The authors are grateful to Dr. Ralph Folsom, Research Triangle Institute, NC, for helpful discussions during early stages of the study, and to Statistics Canada, where Pfeffermann was a research fellow when the article was completed.

[1]Italicized numbers in parentheses cite sources listed in the References section at the end of this article.

[2]Strata consist in general of groups of adjacent counties called Crop Reporting Districts (CDR's). Typically, a State has 8-9 CRD's, with each CRD consisting of 8-10 counties. However, urban-influenced counties have been extracted from CRD's and placed in special strata. In what follows, this definition of strata also defines the survey strata.

estimate these means by the observed sample means, that is, by averaging the farmland values reported by the farmers in the corresponding strata. These strata estimates could then be averaged to produce State and regional estimates, the usual USDA procedure. However, the sample sizes selected by the ALVS within many States are too small to guarantee reliable State estimates, partly because of low response rates. (The effective sample size in some States is less than 40.) For instance, individual State indexes are not constructed for New England (*27*). On the other hand, for selected States where the estimates are more reliable, the National Agricultural Statistics Service (NASS) publishes indexes for strata. Indexes for the United States and 10 major farm production regions may be considered reliable due to sufficiently large samples.

The problem underlying the production of farmland indexes is a typical small-area estimation problem, an issue receiving increased attention in the statistical literature in recent years. The problem of small-area estimation is that because of the small sample size in any given area, the direct survey estimator based only on the sample observed for that area can become very unstable. To overcome that problem, a variety of techniques has been proposed which essentially "borrow strength" from one small area to the next, increasing the precision of the estimators in given small areas.

The data used for these estimators include the observation on the target variable (the ALVS farmers' reports in our study) and the values of regressor variables $x_{1 \ldots} x_k$ representing known small-area characteristics related to the unknown small-area means $\{\theta_i\}$. Denote by $\underset{\sim}{Y}_i$ the vector of observations on the target variable in small area i based on a sample of n_i units. Assuming simple random sampling, it follows that $\overline{Y}_i = \sum_{j=1}^{n_i} Y_{ij} / n_i = \theta_i + \bar{\varepsilon}_i$ where $\bar{\varepsilon}_i$ is the corresponding mean of the error terms $\varepsilon_{ij} = Y_{ij} - \theta_i$, with expectation $E(\bar{\varepsilon}_i) = 0$ and variance $\mathrm{var}(\bar{\varepsilon}_i) = \sigma^2_\varepsilon / n_i$.

When the variances $\sigma^2_\varepsilon / n_i$ are suitably small, the statistician may be content to use the estimates $\overline{Y}_i$, which are basically the estimates currently used by USDA. In view of the small sample sizes (n_i), however, other estimators have to be considered. One natural candidate is the regression estimator $\underset{\sim}{x}'_i\underset{\sim}{b}$ where $\underset{\sim}{b}$ is a vector of estimated regression coefficients based on the individual observations Y_{ij}. The estimator $\underset{\sim}{x}'_i \underset{\sim}{b}$ would be ideal if, for every i, $\theta_i = \underset{\sim}{x}'_i \underset{\sim}{\beta}$ where $\underset{\sim}{\beta}$ represents the "true" unknown regression coefficients. In fact, the estimator $\underset{\sim}{x}'_i\underset{\sim}{b}$ may still be used even when the relationship $\theta_i = \underset{\sim}{x}'_i\underset{\sim}{\beta}$ does not hold provided that the deviations $\{\theta_i - \underset{\sim}{x}'_i\underset{\sim}{\beta}\}$ are sufficiently small.

Often the sample sizes n_i are too small to allow the use of the estimators $\{\overline{Y}_i\}$, and the deviations $\{\theta_i - \underset{\sim}{x}'_i \underset{\sim}{\beta}\}$ are too large to use the estimators $\underset{\sim}{x}'_i\underset{\sim}{b}$. Small-area estimation techniques are essentially a collection of models and inference procedures proposed in the literature to yield estimators that compromise between the estimators $\{\overline{Y}_i\}$ and $\underset{\sim}{x}_i\underset{\sim}{b}$, resulting in increased efficiency.

For example, suppose it can be postulated that $\theta_i = \underset{\sim}{x}'_i \underset{\sim}{\beta} + v_i$, where $E(v_i) = 0$ and $\mathrm{var}(v_i) = \sigma^2_V$. Notice that the deviations v_i are viewed now as random quantities. Under this assumption, the model holding for the original observations can be written as a mixed linear model, $Y_{ij} = \underset{\sim}{x}'_i\underset{\sim}{\beta} + v_i + \varepsilon_{ij}$, where the individual error terms $\{v_i + \varepsilon_{ij}\}$ are now correlated within small areas due to the common effect v_i. This model has been used by Battese, Harter, and Fuller (5) for the estimation of crop areas in counties in Iowa by using satellite data as the regressors. The estimators derived under this model have the general form $\hat{\theta}_i = K_i\overline{Y}_i + (1\text{-}K_i)\, \underset{\sim}{x}'_i\underset{\sim}{b}$, where $K_i = \hat{\sigma}^2_V / \hat{\sigma}^2_V + \hat{\sigma}^2_\varepsilon / n_i)$, and $\hat{\sigma}^2_V$ and $\hat{\sigma}^2_\varepsilon$ define suitable estimators of the unknown variances. The estimator $\hat{\theta}_i$ is a weighted average of the estimators $\overline{Y}_i$ and $\underset{\sim}{x}'_i\underset{\sim}{b}$ with weights that reflect the relative precision of each of the two estimators. In another article, Pfeffermann and Barnard (*17*) synthesize the recent research in small area estimation. In what follows, we refer to that article using the abbreviation P-B.

The model used in the present study extends the Battese-Harter-Fuller model by accounting for both State effects and nested county effects. Let Y_{sci} be the farmland value reported in the ALVS by farmer i residing in county c of State s. Let θ_{sc} stand for the unknown mean market value in county sc. We postulate the following relationship:

$$Y_{sci} = \theta_{sc} + \varepsilon_{sci}; \quad \theta_{sc} = \underset{\sim}{x}'_{sc}\underset{\sim}{\beta} + \alpha_s + \gamma_{sc}, \tag{1}$$

for $s=1 \ldots S$, $c=1 \ldots C(s)$, $i=1 \ldots n_{sc}$,

where $\{\varepsilon_{sci}\}$ are independent errors with zero mean and variance σ^2_e, $\{\alpha_s\}$ represent random State effects with zero mean and variance σ^2_b, and $\{\gamma_{sc}\}$ are random county effects, nested within the State effects with zero mean and variance σ^2_w. We assume that the three random components are mutually independent. S is the total number of States in the study; C(s) is the total number of counties in State s; and n_{sc} is the number of reports in county c of State s.

Equation 1 postulates that the land values reported by farmers residing in the same county, Y_{sci}, are distributed randomly around the true county mean, θ_{sc}. The variation of the county means between counties is modeled as a function of known regressor variables, $\underset{\sim}{x}'_{sc}$, and random State and county effects. The regres-

sor variables represent k county characteristics with typical values represented by $\underset{\sim}{x}'_{sc} = (1, x_{sc1}, \ldots, x_{sck})$ for county sc. (See the next section for the list of regressor variables used in our study.)

The State effects represent any systematic influences on the prices of farmland that are common to all counties in a State, but that are not represented by the regressor variables. State income and property tax laws, State environmental laws, and other regulatory policies that vary by State and restrict farm operation or landownership come to mind. Similarly, the residual county effects represent unique county characteristics that systematically affect the values of farmland, but again, that are not represented by the regressor variables. Examples might be the level of social services, school quality, and other characteristics that affect the quality of life.

Substituting the right-side equation of (1) into the left-side equation gives the mixed linear (components of variance) model representation:

$$Y_{sci} = \underset{\sim}{x}'_{sc}\underset{\sim}{\beta} + \alpha_s + \gamma_{sc} + \varepsilon_{sci}, \tag{2}$$

which implies:

$$VAR(Y_{sci}) = \sigma_b^2 + \sigma_w^2 + \sigma_e^2 ;$$

$$COV(Y_{sci}, Y_{sci^*}) = \sigma_b^2 + \sigma_w^2 , \; i \neq i^*;$$

$$COV(Y_{sci}, Y_{sc^*i^*}) = \sigma_b^2 , \; c \neq c^*;$$

$$COV(Y_{sci}, Y_{s^*c^*i^*}) = 0 , \; s \neq s^*. \tag{3}$$

Thus, the model states that values reported by farmers residing in the same county are correlated, as are values reported by farmers residing in the same State but in different counties.

The actual application of the model requires as a first step the identification of available data sources to be used as potential regressor variables. We discuss this issue in the next section. The formulas of the predictor of the county and State means, as obtained under the present model, are given later. Note in this regard, that since the county and State means are considered as random under the model, we adopt hereafter the conventional statistical terminology and refer to the assessment of these means as "prediction" rather than "estimation."

Available Data: Sources and Definitions

Identification of factors that affect farmland values and statistical measurement of their importance has been the objective of studies for more than 60 years, spawning an extensive literature within the agricultural economics profession. Reynolds (*18*), for instance, cites a partial list of more than 60 empirical studies. The purpose of most of these studies was to discover the determinants of variation in farmland values and estimate the parameters associated with those factors. Empirically, the general procedure is to regress observations on farmland values against corresponding observations on a set of independent variables representing variation in productivity or income, location relative to markets and services, and nonagricultural influences.

In contrast, the purpose of our procedure is to identify regressor variables, which, when used jointly, can best predict the county and State mean farmland values without worrying about causal relationships and substantive interpretations. Nevertheless, previous models fitted to farmland value data provide a natural basis to guide the selection of factors to include in our model. The other obvious consideration in the preliminary selection of such regressor variables is data availability, which we describe next. We then specify the variables considered in our study.

Sources

The U.S. Department of Commerce (USDC) is a major source of county-level information that can be used as additional information to model the variation of farmland values. The Census of Agriculture, generally conducted every 5 years, provides a wide array of agriculturally related information, including acres of land in farms, numbers of farms, crop acres harvested, quantities of crops and livestock produced (sold), market values of crops and livestock sold, and days of off-farm work (*28*). Indeed, much of the literature involving cross-sectional analyses of aggregate farmland values has used county, State, or national estimates provided by the Census of Agriculture. In addition, the decennial Census of Population collects information on rural and urban population. The Bureau of Economic Analysis, through its Regional Economic Information System, provides annual data on local area employment and personal income, by Standard Industrial Category (SIC). The data series available include mineral income, net farm income, and off-farm employment.

A notable feature of the procedure presented in this article is its ability to include *alternative assessments of farmland value* among the regressor variables. USDA, itself, collects farmland value information

from three other independent sources, which provides direct assessments of county farmland values.[3] Of particular interest is a set of data collected annually by the Agricultural Stabilization and Conservation Service (ASCS) from each of its County Executive Directors (CED's). This variable represents the opinions of ASCS county executive directors, one per county, concerning the average value of nonirrigated cropland in their county. Most CED's consult with farmers, lenders, and other real estate professionals before forming their opinions. The data represent the opinion of each CED regarding average values of farmland in each county whose ASCS program they administer. While there are some small differences in definition, these data provide an independent assessment of farmland values (*1,8*) and are particularly valuable because reports are received from virtually every agricultural county in the United States. (See also the last two paragraphs at the end of this section.)

Other sources of data on farmland values include the Farmland Market Survey, which obtains both sales data on individual tracts and opinion data on county farmland values (*27*). The opinion data from this survey are similar in form to the ASCS data but lack estimates for all counties. The USDA Farm Costs and Returns Survey (FCRS) (*25*) and the Census of Agriculture (*3,28*) are sources of information on the value of farmland and buildings. The Census of Agriculture provides data for every agricultural county but only at 5-year intervals. Data from the FCRS, though available annually, are not available for every county.

Variables Considered for Analysis

In studies cited by Reynolds (*18*) and Reynolds and Comer (*19*), many variables were found to be important determinants of farmland values. Particular data used and specific results obtained have depended upon data availability and the level of aggregation employed. Data used have varied from microdata on sales of individual tracts to aggregate data collected on a State, regional, or national basis (*2*). In the selection and specification of variables for our model, we relied mostly on cross-sectional studies that used county data. The variables chosen are general in the sense that they could be used in the analysis of nonirrigated cropland in most regions of the United States. Our initial model included 10 variables: eight that represent various aspects of agricultural productivity and urban influence, one that represents mining activity, and one that is the independent assessment of nonirrigated cropland value from the CED's of ASCS. Brief descriptions of the variables selected, their sources, and abbreviated names are provided in table 1, with more detailed explanations following.

Various measures have been used to represent agricultural productivity and the overall economic potential of farmland (*4,6,9,15,16,21,22*). In our model, PCTFARM and PCTGRAZ were included in the model to represent the basic suitability of the land for crops, which depends on soil, climate, topography, and other factors. A larger percentage of farmland generally indicates higher average productivity, while larger percentages of grazing land indicate lower average productivity. Variation in overall economic potential of land for agricultural use is measured by FARMINCOME and CROPSVALU. Larger net farm incomes and gross crop receipts per acre imply more productive cropland. Although these variables are crude measures individually, taken together they serve as proxies for the agricultural value of farmland.

A similar variable, SPECLTYVALU, was included to capture the contribution of high-valued specialty crops, including vegetables, fruits, berries, nuts, and greenhouse products. Such variables measure differences in land use intensity. The importance of specialty crops as a determinant of land values is demonstrated by Reynolds and Tseng (*21*) in a study of Florida counties.

Size of tract has been demonstrated to be an important explanatory variable in models designed to explain farmland values (*6,7,12,15,16,21,22,23*). Our variable, FARMSIZE, serves as a proxy for size of tract sold. Value per acre declines as tract size increases, *ceteris paribus*.

Measures of urban influence, including population, population density, and extent of off-farm employment, have often been found to have large and statistically significant effects on farmland value (*6,7,9,12,13,15,16,20,21,22,23*). In our study, these nonagricultural influences are represented by POPULATN and NUMOFFFARM. Larger urban populations imply increased demand for farmland for rural residences. More off-farm employment opportunities imply increased potential for part-time and hobby farms. Nonagricultural uses often can outbid agriculture for use of urban-influenced farmland.

MINEINCOME is another variable related to nonagricultural influences, included principally to capture the effect that mineral rights may have on the sale price of individual parcels. When opinions of value are formed on the basis of reported farmland sale prices, a portion of the value of mineral rights

[3] All four USDA sources collect data during January-February of each year.

Table 1—Variables used in the empirical study

Abbreviated name	Description	Source
CED	County executive directors' opinions of mean county farmland values	ASCS questionnaire[1]
PCTFARM	Acres of farmland as a percentage of county land area	Census of Agriculture[2]
POPULATN	Urban population per acre of total cropland	Census of Population[2] Census of Agriculture[2]
CROPSVALU	Market value of crops sold per acre of total cropland	Census of Agriculture
FARMINCOME	Net farm income per acre of land in farm	Local area personal income[3]
FARMSIZE	Average number of acres per farm	Census of Agriculture
PCTGRAZ	Acres of grazing land as a percentage of land in farms	Census of Agriculture
SPECLTYVALU	Market value of specialty crops per acre of total cropland	Census of Agriculture
NUMOFFFARM	Number of farm operators who worked at least some days off the farm	Local area employment[3]
MINEINCOME	County income from mining	Local area personal income[3]

[1]Agricultural Stabilization and Conservation Service, USDA.
[2]Bureau of the Census, USDC.
[3]Bureau of Economic Analysis, USDC.

may be incorporated. This factor may positively affect cropland values in areas with substantial oil, gas, and coal development.

The final variable included in our initial model was the average value of nonirrigated cropland from ASCS described earlier, labeled CED in table 1. The inclusion of the CED variable as one of the regressor variables raises two interesting questions. The first question refers to the different roles assigned to the CED and the ALVS measurements, namely, one variable being specified as an independent variable and the other as the dependent variable, despite the fact that both variables measure essentially the same phenomenon. Our consideration in including the CED variable as the regressor variable was that this variable, unlike the ALVS, is measured in every county and can be used in the model without missing observations. Also, in contrast to the ALVS estimates, whose precision depends on the realized sample sizes, which differ from one county to the other, the CED variable uses the same sort of information in every county. Theoretically, a better way to include these variables in the model would have been to specify both of them as dependent correlated variables. Notice, however, that this multivariate framework is much more complicated computationally, whereas the gains in terms of the efficiency of the resulting predictors would generally be low considering that both the univariate and the multivariate models exploit the

same amount of information. If the joint distribution of the two estimators can be assumed to be bivariate normal, then the structure of the predictors as obtained under the two models is similar (even though not the same).

The other question applies to the interpretability of the model. In some sense, the CED variable encompasses and measures the interaction of all the other regressor variables included in the model and as such, the model has no longer a substantive causal interpretation. We re-emphasize, however, that the purpose of the analysis is the prediction of the county and State means. Thus, variables have been included in the model based on their prediction power and not with respect to their substantive interpretation, an important factor when analyzing the results of this study.

Computation of the Predictors and Prediction MSE's

In this section, we outline the major stages in fitting the model defined by equations 2 and 3 to the actual data. A more technical and comprehensive discussion can be found in the P-B article. We assume a given set of regressor variables with typical values $\underset{\sim}{X}'_{sc} = (1, x_{sc1} \ldots x_{sck})$ corresponding to county sc.

Presentation of the Model in Matrix Notation

Let $\underset{\sim}{Y}_{sc}$ represent the vector of observed values in county sc, and let $\underset{\sim}{Y}'_s = (\underset{\sim}{Y}'_{s1} \ldots \underset{\sim}{Y}'_{sC(s)})$ define the vector of observations in State s so that $\underset{\sim}{Y}' = (\underset{\sim}{Y}'_1 \ldots \underset{\sim}{Y}'_S)$ defines the entire vector of reported land values. A similar notation is used for the residuals $\{\varepsilon_{sci}\}$. We denote by $\underset{\sim}{\alpha}' = (\alpha_1 \ldots \alpha_s)$ the vector of State effects and by $\underset{\sim}{\gamma}' = (\gamma_{11} \ldots \gamma_{1C(1)}, \ldots, \gamma_{S1} \ldots \gamma_{SC(S)})$ the vector of nested county effects of order $T_\gamma = \sum_{s=1}^{S} C(s)$.

Using the symbol $\otimes$ to define the Kronecker product, $n_s = \sum_{c=1}^{C(s)} n_{sc}$ to represent the number of observations in State s, and $\underset{\sim}{1}'_m$ to define in general a $1 * m$ vector of ones, the model defined by equation 2 can be written compactly as:

$$\underset{\sim}{Y} = X\underset{\sim}{\beta} + Z_b\underset{\sim}{\alpha} + Z_w\underset{\sim}{\gamma} + \underset{\sim}{\varepsilon} = X\underset{\sim}{\beta} + \underset{\sim}{\mu}, \tag{4}$$

where $X' =$

$$[\underset{\sim}{1}'_{n_{11}} \otimes \underset{\sim}{x}_{11}, \underset{\sim}{1}'_{n_{12}} \otimes \underset{\sim}{x}_{12}, \ldots, \underset{\sim}{1}_{n_{1C(1)}} \otimes \underset{\sim}{x}_{1C(1)} \ldots$$

$$\ldots \underset{\sim}{1}'_{n_{S1}} \otimes \underset{\sim}{x}_{S1}, \ldots, \underset{\sim}{1}'_{n_{SC(S)}} \otimes \underset{\sim}{x}_{SC(S)}]$$

$$Z_b = \begin{bmatrix} \underset{\sim}{1}_{n_1} & & & \\ & \ddots & & \\ & & & \underset{\sim}{1}_{n_S} \end{bmatrix}$$

$$Z_w = \begin{bmatrix} \underset{\sim}{1}_{n_{11}} & & & \\ & \ddots & & \\ & & & \underset{\sim}{1}_{n_{SC(S)}} \end{bmatrix}$$

The vector $\underset{\sim}{\mu}$ satisfies:

$E(\underset{\sim}{\mu}) = \underset{\sim}{0}$;

$$E(\underset{\sim}{\mu}\underset{\sim}{\mu}') = \sigma_b^2 Z_b Z'_b + \sigma_W^2 Z_W Z'_W + \sigma_e^2 I_n, \tag{5}$$

where $n = \sum_{s=1}^{S} n_s$ and I_n is the identity matrix of order n.

Optimal Predictors of County and Strata Means Assuming Known Variances

The optimal predictors of the county and Strata means are obtained in a straightforward manner from the optimal predictor of the vector $\underset{\sim}{\lambda}' = (\underset{\sim}{\beta}', \underset{\sim}{\alpha}', \underset{\sim}{\gamma}')$. One way to derive the optimal predictor $\hat{\underset{\sim}{\lambda}}$ and the associated variance-covariance (V-C) matrix of the prediction errors $(\hat{\underset{\sim}{\lambda}} - \underset{\sim}{\lambda})$ is to compute $\hat{\underset{\sim}{\lambda}}$ as the generalized least squares (GLS) solution of the regression model:

$$\underset{\sim}{Y}^0 = \begin{bmatrix} \underset{\sim}{Y} \\ \\ \underset{\sim}{0}_r \end{bmatrix} = \begin{bmatrix} X, Z_b, & Z_w \\ \\ {}_r0_{k+1}, & -I_r \end{bmatrix} \begin{bmatrix} \underset{\sim}{\beta} \\ \underset{\sim}{\alpha} \\ \underset{\sim}{\gamma} \end{bmatrix} + \begin{bmatrix} \underset{\sim}{\varepsilon} \\ \underset{\sim}{\alpha} \\ \underset{\sim}{\gamma} \end{bmatrix} = X^0\underset{\sim}{\lambda} + \underset{\sim}{\varepsilon}^0, \tag{6}$$

where $\underset{\sim}{0}_r$ and ${}_r0_{k+1}$ define correspondingly a column null vector of order $r = S+T_\gamma$ and a null matrix of order $r * (k+1)$ (*10*). The error vector $\underset{\sim}{\varepsilon}^0$ satisfies $E(\underset{\sim}{\varepsilon}^0) = \underset{\sim}{0}$, $E(\underset{\sim}{\varepsilon}^0\underset{\sim}{\varepsilon}^{0\prime}) = V = \text{Diag}[\sigma_e^2\underset{\sim}{1}'_n, \sigma_b^2\underset{\sim}{1}'_S, \sigma_W^2\underset{\sim}{1}'_{T_\gamma}]$. The GLS estimator of $\underset{\sim}{\lambda}$ is:

$$\hat{\underset{\sim}{\lambda}} = (X^{0\prime}V^{-1}X^0)^{-1} X^{0\prime}V^{-1}\underset{\sim}{Y}^0. \tag{7}$$

Notice that X^o is of full rank (assuming X is of full rank), which guarantees a unique solution. The V-C matrix of the prediction errors has the common form

$$VAR(\hat{\lambda} - \lambda) = E(\hat{\lambda} - \lambda)(\hat{\lambda} - \lambda)' = (X^{o\prime}V^{-1}X^o)^{-1}, \qquad (8)$$

where the block matrix consisting of the first (k+1) rows and columns of $(X^{o\prime}V^{-1}X^o)^{-1}$ is the V-C matrix of the GLS estimator $\hat{\beta}$ of β.[4]

The optimal predictors of the county means and the corresponding prediction variances are obtained from $\hat{\lambda}$ and $VAR(\hat{\lambda} - \lambda)$ as:

$$\hat{\theta}_{sc} = x'_{sc}\hat{\beta} + \hat{\alpha}_s + \hat{\gamma}_{sc} = h'_{sc}\hat{\lambda}, \qquad (9)$$

$$E(\hat{\theta}_{sc} - \theta_{sc})^2 = h'_{sc}(X^{o\prime}V^{-1}X^o)^{-1}h_{sc}, \qquad (10)$$

where $h'_{sc} = (x'_{sc}, q'_{sc})$ and q'_{sc} is a row vector of length $(S+T_\gamma)$ with 1's in positions s and $(\sum_{t=0}^{s-1} C(t) + c)$ and zeros elsewhere; $[C(0) \equiv 0]$.

The mean farmland values of the survey strata are obtained from the county means as:

$$\theta_{sh} = \sum_{sc\in sh} a_{sc}\theta_{sc} / \sum_{sc\in sh} a_{sc} = \sum_{sc\in sh} \tilde{a}_{sc}\theta_{sc}, \qquad (11)$$

where a_{sc} is the total acreage of the particular type of farmland in county sc, $\tilde{a}_{sc}$ is the proportion of acreage of that particular type of farmland in stratum sh that is found in county sc, and the summation is over counties in State s belonging to stratum h. Since θ_{sh} is a linear combination of the county means, it follows that the optimal predictors of the survey strata means are:

$$\hat{\theta}_{sh} = \sum_{sc\in sh} \tilde{a}_{sc}\hat{\theta}_{sc} = \hat{\alpha}_s + \sum_{sc\in sh} \tilde{a}_{sc}(x'_{sc}\hat{\beta} + \hat{\gamma}_{sc}) = f'_{sh}\hat{\lambda}, \qquad (12)$$

where $f'_{sh} = (\sum_{sc\in sh} \tilde{a}_{sc}x'_{sc}, r'_{sh})$ and r'_{sh} is a row vector of length $(S + T_\gamma)$ with one in position s, $\tilde{a}_{sh1} \ldots \tilde{a}_{shC(sh)}$ in the positions corresponding to counties included in stratum sh, and zeros elsewhere. C(sh) is the number of counties included in stratum h of State s. For example, s=1, if S=5, and h=1, then $r'_{11} = (1,0,0,0,0,|\,\tilde{a}_{111} \ldots \tilde{a}_{1C(11)},0,\ldots 0|\,0,\ldots 0)$. The prediction variance of $\hat{\theta}_{sh}$ is:

$$E(\hat{\theta}_{sh} - \theta_{sh})^2 = f'_{sh}(X^{o\prime}V^{-1}X^o)^{-1}f_{sh}. \qquad (13)$$

The use of equations 9-13 assumes that the sample includes farmers from every county. P-B gives the appropriate formulas for the case where some of the counties are not represented in the sample. The optimal predictors of the State means can be obtained in similar fashion.

Variance Estimation

The discussion to this point assumes known variances. In practice, the variances have to be estimated from the sample. P-B discusses the practical aspects of estimating the unknown variances by maximum likelihood methods assuming that the model random disturbances have a normal distribution. They illustrate that the variance estimates can be obtained by iterating between the procedures "REG" and "VARCOMP" in SAS.

Substituting the sample estimates for the true variances in the formulas for $\hat{\theta}_{sc}$ and $\hat{\theta}_{sh}$ gives the corresponding empirical predictors of the county and strata means. Performing a similar substitution in the formulas of the V-C matrices yields, in the case of large samples, the V-C matrices of the empirical predictors. These matrices have to be modified in the case of small sample sizes in order to account for the extra variability induced by the need to estimate the unknown variances. See, for example, Kackar and Harville (*11*).

Application of the Model

The model defined in the previous section was applied to data collected by the ALVS. The purpose of this analysis was twofold: to test the suitability of the model to the land values data, and to compare the performance of the model-dependent predictors with the performance of other possible predictors (estimators), including the survey estimator used by USDA.

The USDA Survey Estimator

The ALVS is an opinion survey of farmers and ranchers. Participants in the survey are selected by a stratified simple random design, carried out separately within each of the States, with a 20-percent sample rotation from one year to the next. The questionnaire asks for information on average market value per acre of irrigated and nonirrigated cropland, grazing land, and woodland. The values reported by the farmers are averaged first within strata and then over the strata within States to yield estimates of

[4] An important advantage of expressing $\hat{\lambda}$ as the GLS solution of the regression model (equation 6) is that the predictor and the prediction V-C matrix can be computed using any computer software for weighted regression with Y^0 as the dependent variable, X^o as the design matrix, and $w' = (\tilde{1}'_{n+r})V^{-1}$ as the vector of weights.

State average market value, by type of farmland (*14*). Until 1989, the averages within strata were simple means, while the averages of the strata means were weighted averages, the weights being relative to the total acreage of the particular type of farmland in the given strata. ERS changed its procedure in 1989, and the strata estimates are now weighted averages of county means. Acreages come directly from, or are derived from, the latest Census of Agriculture (*28*).

Application of the Model to Corn Belt Data

The survey data analyzed in this study are the values of nonirrigated cropland in the Corn Belt States as collected in the 1984 ALVS.[5] Nonirrigated cropland constitutes the major land use in the region. The data consist of 871 farmers' reports representing 5 States (Indiana, Illinois, Iowa, Missouri, and Ohio), 43 strata, and 251 counties. We excluded from the analysis the strata formed for the urban-influenced counties (see footnote 2) since the farmland values in these strata behave very differently from the values in the other strata, thus requiring extra treatment.[6] In urban-influenced counties, particularly those that are part of large metropolitan areas, farmland values are higher and have larger variances than counties in more rural areas. The mean and variances of farmland values in the excluded strata are 38 percent and 339 percent higher than in the remaining strata, respectively. Farmland values in urban-influenced counties exhibit little relationship to the agricultural characteristics that determine farmland values in rural counties, suggesting the need for alternative model specification. Although the current model does contain a proxy for urban influence (the POPULATN variable), county-level population cannot fully account for the influence of large multicounty metropolitan areas. Distance from the center of the county to the center of the nearest major metropolitan area might more accurately account for the variation in the excluded strata. Distance measures have been used in previous studies with good success. Such measures are not available from published sources, but future work should involve the development of such data.

The 10 variables listed in table 1, plus an added intercept, formed the initial X matrix for the model (equation 4), while the dollar per acre values reported in the ALVS constituted the Y vector. The model was estimated based on the entire data set. The significance of the β coefficients was tested by using the Wald statistic (*29*). The six variables listed in the lower part of the table and the intercept variable were jointly insignificant in the presence of the other four variables. (As discussed before, the emphasis in the present study is on prediction rather than on interpretation, so we chose to include variables with significant predictive power rather than variables necessarily having substantive interpretation.) Consequently, the non-significant variables were excluded from the model and were not considered in the rest of the analysis. (The Wald statistic for testing a hypothesis of the form H_0: $C\underset{\sim}{\beta} = 0$, where C is r * (k+1), is $W = (C\hat{\underset{\sim}{\beta}})'$ [C VAR($\hat{\underset{\sim}{\beta}}$) C']$^{-1}C\hat{\underset{\sim}{\beta}}$, and it has an asymptotic chi-square distribution with r degrees of freedom under H_0. The value observed when testing the joint significance of the seven variables was W=4.15, which was well below the customary critical values of that $\chi^2_{(7)}$ distribution.)

Table 2 shows the four significant regression coefficients (first four elements of the empirical predictor $\underset{\sim}{\hat{\lambda}}$) along with their estimated standard errors, the variance components estimates, and twice the log of the likelihood ratio test statistic (log LRT) used for testing significance. These test values indicate highly significant variance component estimates as can be seen by comparing the test values to critical values of the $\chi^2_{(1)}$ distribution. The test results should be interpreted with caution, since the postulated chi-square distribution is a large sample property, whereas the data represent only five groups.

Table 2 reveals the highly significant nature of the CED variable, which is by far the most important predictive variable. To illustrate the importance of this variable, we conducted the following simple analysis, using ordinary least squares regression (OLS). An equation containing only the CED variable and an intercept was estimated and compared with an equation containing the four significant variables and an intercept. The regression sum of squares for the CED-only equation amounted to 96 percent of the regression sum of squares for the latter equation. Dropping the CED variable and estimating an equation containing only an intercept and the other three significant variables results in a 30-percent reduction in the regression sum of squares.

The dominant predictive power of the CED variable (available from ASCS data) is especially important because the information it contains is updated annually and in the same time period as the ALVS. This contrasts with the 5-year periodicity of information from the Census of Agriculture.

[5]Restriction of the analysis to the Corn Belt was mainly for technical reasons, but this region, nevertheless, sufficiently illustrates the important features of the proposed procedures.

[6]The Corn Belt consists of 495 counties: 49 are part of the excluded urban-influenced strata and 195 had no observations in the ALVS.

Table 2—Significant regression coefficients and variance components

Item	CED	PCTFARM	POPULATN	CROPSVALU
Regression coefficients	0.59	663.8	357.4	811.1
Standard errors	.05	118.9	148.1	498.3

Variance components	Significance tests	2 logLRT
Between States—$\sigma_b^2 = 24{,}337$	$H_0{:}\sigma_b^2 = 0$	60.2
Between counties—$\sigma_w^2 = 24{,}157$	$H_0{:}\sigma_w^2 = 0$	31.6
Residual—$\sigma\pi_e^2 = 174{,}940$		

Testing the Performance of the Model

To assess the performance of the model in predicting the unknown strata and State means, we performed a cross-validation study by which the model-based predictor and other estimators were calculated based on one part of the sample (the estimation part). The performance of the predictor and estimators has been evaluated based on their quality in predicting the data included in the complementary part (the validation part). This method differs from the direct analysis of all the data reported in the P-B article, with the advantage that the assessment and comparison of the various estimators and predictors are less tied to a particular model. The results obtained from the study, however, refer to the sample sizes of the partitioned data sets and not to the sample sizes of the combined sample, which are the actual sample sizes of the ALVS.

We split the sample between counties within strata. About half the counties of each stratum were allocated to the estimation part and the other half to the validation part. We employed a simple random sampling design for the splitting algorithm.

We evaluated the performance of four predictors of the survey strata means by computing the prediction bias and root mean square error (RMSE) of the predictors and averaging the results within States by using the relative strata acreages as weights. The strata-based analysis enables a comparison with the survey estimator used until 1989, which is defined by USDA as an unweighted average at the strata level.[7] Thus, let $\tilde{M}_{sh}$ represent any one of the four predictors and $M_{sh} = \sum_{sc \in V_{sh}} a_{sc} \overline{Y}_{sc} / \sum_{sc \in V_{sh}} a_{sc}$ define the mean for farmers included in the validation part of stratum h in State s, where a_{sc} is the acreage of nonirrigated cropland in county sc and $\overline{Y}_{sc}$ is the sample mean of observations in county sc. As such, the prediction BIAS and RMSE are represented by:

$$\mathrm{BIAS}_s(\tilde{M}_{sh}) = \left(\sum_h a_{sh}(\tilde{M}_{sh} - M_{sh})\right) / \sum_h a_{sh}, \tag{14}$$

and

$$\mathrm{RMSE}_s(\tilde{M}_{sh}) = \left(\sum_h a_{sh}(\tilde{M}_{sh} - M_{sh})^2 / \sum_h a_{sh}\right)^{1/2}, \tag{15}$$

where a_{sh} is the acreage of nonirrigated cropland in stratum sh and the summation $\sum_h$ is over all the strata included in State s.

Using the prediction bias (equation 14) and RMSE (equation 15) as criteria, we compare the performance of the following predictors of the strata means:

A. The USDA survey estimates, $\hat{M}_{sh}$, which were defined as:

$$\hat{M}_{sh} = \sum_{sc \in E_{sh}} n_{sc} \overline{Y}_{sc} / \sum_{sc \in E_{sh}} n_{sc}, \tag{16}$$

where the summation is over counties from stratum h in State s included in the estimation part.

[7] Our cross-validation study was initially designed to evaluate the pre-1989 USDA estimator. Since we are trying to predict strata means over counties included only in the validation part, there was no apparent reason to prefer the new USDA estimator over the old USDA estimator. Supplemental analysis indicates that a comparison between our estimator and either of the two USDA estimators is essentially independent of the weighting procedure used. In the augmented analysis, we considered a second split, which allocated approximately half of the farmers of each county to the estimation part and the rest of the farmers to the validation part. To reflect more closely the new procedure used by USDA in 1989, we weighted our county predictions by county acreages. The results obtained for that second split are generally consistent with results reported here.

B. The optimal predictors, $(\hat{\theta}^E_{sh})$, where the superscript "E" added to the previous notation is used to emphasize that the predictors have been calculated based on the estimation part and that the unknown variances have been replaced by the sample estimates. The optimal strata means predictors are defined as:

$$\hat{\theta}^E_{sc} = \sum_{sc \in V_{sh}} a_{sc}\, \hat{\theta}^E_{sc} \Big/ \sum_{sc \in V_{sh}} a_{sc}, \tag{17}$$

where

$$\hat{\theta}^E_{sc} = \underset{\sim}{x}'_{sc}\underset{\sim}{\hat{\beta}} + \hat{\alpha}_s. \tag{18}$$

The county effects, $\hat{\gamma}_{sc}$, are estimated as zero because the sample was split between counties, so that counties sc selected for the validation part are not represented in the sample.

C. The synthetic regression estimators, $\hat{R}_{sh}$, which are calculated as weighted averages of the county regression estimators, $\hat{R}_{sc} = \underset{\sim}{x}'_{sc}\underset{\sim}{\hat{\beta}}$, where $\underset{\sim}{\hat{\beta}}$ is the optimal maximum likelihood estimator (mle) of $\underset{\sim}{\beta}$ and the weighting procedure used is the same as that defined above for the optimal estimators.

D. The synthetic regression estimators, $\hat{R}^{ols}_{sh}$, which are calculated in the same way as the estimator $\hat{R}_{sh}$ except that $\underset{\sim}{\beta}$ is estimated using ordinary least squares.

The synthetic regression estimators, $\hat{R}_{sh}$ and $\hat{R}^{ols}_{sh}$, represent alternative estimators that also incorporate the county-specific information. The estimator $\hat{R}_{sh}$ accounts for the correlations between the various farmers' opinions which result from the common county and State effects (see equation 3). Specific estimates of the State and county effects, however, are not incorporated into this estimator. The OLS estimator, on the other hand, ignores State and county effects altogether.

Table 3 gives the prediction bias and RMSE of the various predictors separately for each State. Also shown are the target weighted averages of the strata means in the validation part defined as $M_s = \sum_h a_{sh}M_{sh} \Big/ \sum_h a_{sh}$.[8]

[8]A robust predictor, incorporating a restriction to assure that the mean farmland value predicted under the model for the entire group of States will equal the survey estimator of that same mean, is derived in P-B. The bias and RMSE of the robust predictor came out very similar to those of the optimal predictor. This outcome can be considered indicative of the adequacy of the model.

The main conclusion to note from the table is that the use of the alternative data sources improves the prediction of the farmland values. The improvement is evident at the State level as revealed by comparing the prediction biases of the optimal predictor and the USDA survey estimator. The prediction bias of the optimal predictor is substantially lower in four of the five States. Among the three predictors using the additional information, the optimal predictor is clearly the most accurate, demonstrating the benefit of accounting for State and county effects in the form of a variance components model. The two synthetic regression estimators show improvement relative to the USDA survey estimator in the prediction of the State means in two States, but the estimators actually perform less well than the USDA survey estimators in the other three States, particularly in Missouri where they miss by a wide margin.

The RMSE of the optimal estimator is lower than the RMSE of the USDA survey estimator in three States. The reduction amounts to approximately 50 percent in two of those States. The RMSE of the optimal estimator in the remaining two States is only slightly larger than for the USDA survey estimator. The two synthetic estimators also show a reduction in the RMSE relative to the USDA survey estimator in three States, but the reduction is less pronounced than for the optimal estimator. For Missouri, the RMSE's of the synthetic estimators are considerably larger than the RMSE of the USDA survey estimator.

The use of the additional information not only improves upon the USDA survey estimators in terms of point predictions but also provides a basis for probabilistic inference. Table 4 contains the 95-percent prediction intervals for the validation State means. The prediction intervals are of the form $\hat{\theta}^E_s \pm Z_{\alpha/2}[\widehat{VAR}(\hat{\theta}^E_s - M_s)]^{1/2}$ where the model-dependent estimates of the prediction variances are used in the calculation. The notable result from table 4 is that the validation mean is within the prediction interval. The bias $(\hat{\theta}^E_s - M_s)$ is less than $1.96[\widehat{VAR}(\hat{\theta}^E_s - M_s)]^{1/2}$ in all five States, indicating the insignificance of the prediction bias at the 5-percent level.

Conclusions and Model Extension

The results of the empirical study indicate that the use of alternative data sources improves the precision of mean farmland value estimates. Consideration of

Table 3—Bias and root mean square errors of strata means predictors

Item	Criteria	Indiana	Illinois	Iowa	Missouri	Ohio
				Dollars per acre		
State means (validation)		1,958	1,689	1,663	822	1,484
Predictors:						
$\hat{M}_{sh}$	BIAS	-116.8	49.9	73.5	-25.5	-114.5
	RMSE	378.7	278.0	190.5	128.7	331.7
$\hat{\theta}^{E}_{sh}$	BIAS	-31.4	-4.1	23.5	83.2	-67.5
	RMSE	182.6	214.5	199.0	141.3	184.1
$\hat{R}_{sh}$	BIAS	-187.5	-93.0	-3.1	268.3	19.0
	RMSE	259.8	233.8	197.6	291.6	172.3
$\hat{R}^{ols}_{sh}$	BIAS	-137.2	-60.1	21.4	202.9	19.9
	RMSE	220.4	242.0	207.5	235.1	180.3

Table 4—Confidence intervals for the validation State means

State	Upper limit	Validation mean	Lower limit
		Dollars per acre	
Indiana	2,073	1,958	1,779
Illinois	1,879	1,689	1,507
Iowa	1,815	1,663	1,557
Missouri	1,074	822	736
Ohio	1,580	1,484	1,254

State and county effects in the form of a nested, variance-components model adds to the precision of the assessments. The computations involved in the application of the procedure can be performed using available statistical software. In addition, the model provides a satisfactory basis for probabilistic inference.

Although the study demonstrated the potential for predictors derived under the model to improve substantially upon the estimators used by USDA, the results strictly apply only to a major land use (nonirrigated cropland) in a very homogeneous farm production region (the Corn Belt). The procedure's ability to produce improvements for irrigated cropland, grazing land, and woodland in more heterogeneous regions is yet to be tested. A full evaluation would also involve extension of the model to include more States in the analysis and the consideration of additional regressor variables. The inclusion of more States will provide more stable estimates for the variance components and, hence, better predictors of State and strata farmland values. Notice in this respect that it is unnecessary to assume the same regression coefficients for all regions. By appropriate definition of the design X-matrix, different vectors of coefficients can be postulated for different regions.

Consideration of additional regressor variables may improve the predictions. Variables that jointly account for both population of major metropolitan areas and county location relative to those areas may be especially helpful. Such variables, which represent access to social services, recreational facilities, and other quality-of-life conditions, may be most useful in modeling farmland values in the urban-influenced strata that were excluded from this study.

As a final note, we point out the potential applicability of this procedure to a wide variety of data obtained from surveys conducted by ERS and NASS. With appropriate modification, the procedures could be applied, for example, to farmland value data obtained in the FCRS.

References

1. Barnard, C.H. *U.S. Farmland Values, April 1982: An Experimental Survey of ASCS County Executive Directors.* Staff Report No. AGES830706, U.S. Dept. Agr., Econ. Res. Serv., Sept. 1983.

2. Barnard, C.H. "Data Sources for Farm Real Estate Market Studies," *The Farm Real Estate Market.* SRDC Publication No. 83, Mississippi State Univ., 1986, pp. 15-28.

3. Barnard, C.H., and J. Jones. *Farm Real Estate Values in the United States by Counties, 1850-1982.* SB-751, U.S. Dept. Agr., Econ. Res. Serv., Mar. 1987.

4. Barrows, R.L., and R.W. Dunford. *Agricultural, Urban, and Recreational Determinants of Farmland Values in Wisconsin.* University of Wisconsin, Madison, College of Agricultural and Life Science, Research Bulletin R2764, 1976.

5. Battese, G.E., R.M. Harter, and W.A. Fuller. "An Error Components Model for Prediction of County Crop Acres Using Survey and Satellite Data," *Journal of the American Statistical Association.* Vol. 83, 1988, pp. 28-36.

6. Clifton, I.D. "Using Automatic Interaction Detection (AID) Model to Obtain Homogeneous Classifications of Farmland Markets," *Agricultural Economics Research.* Vol. 27, 1975, pp. 93-100.

7. Greene, C., and C.H. Barnard. "Agricultural Land Markets in U.S. Urban-Fringe Counties," *Landscape and Urban Planning.* Vol. 16, 1988, pp. 293-301.

8. Greene, C., and C.H. Barnard. *U.S. Farmland Values, 1982-84: A Comparison of Experimental and Traditional Data.* Staff Report No. AGES850304, U.S. Dept. Agr., Econ. Res. Serv., July 1985.

9. Hammill, A.E. "Variables Related to Farm Real Estate Values in Minnesota Counties," *Agricultural Economics Research.* Vol. 21, 1969.

10. Harville, D.A. "Maximum Likelihood Approaches to Variance Component Estimation and to Related Problems," *Journal of the American Statistical Association.* Vol. 72, 1977, pp. 320-40.

11. Kackar, R.N., and D.A. Harville. "Approximations for Standard Errors of Estimators of Fixed and Random Effects in Mixed Linear Models," *Journal of the American Statistical Association.* Vol. 79, 1984, pp. 853-62.

12. Morris, D.E. "Farmland Values and Urbanization," *Agricultural Economics Research.* Vol. 30, 1978, pp. 44-47.

13. Norberg, R.P., and J.E. Reynolds. "Population Pressure and Farmland Values in Florida," *Soil and Crop Science Society of Florida: Proceedings.* Vol. 41, 1982, pp. 118-21.

14. North Central Regional Committee on Land Values. *Ongoing Farmland Market Research: A Handbook.* North Central Regional Research Publication No. 306. University of Nebraska, Lincoln, Sept. 1985.

15. Pasour, E.C., Jr. "Real Property Taxes and Farm Real Estate Values: Incidence and Implications," *American Journal of Agricultural Economics.* Vol. 55, 1973, pp. 549-56.

16. _____. "The Capitalization of Real Property Taxes Levied on Farm Real Estate," *American Journal of Agricultural Economics.* Vol. 57, 1975, pp. 539-48.

17. Pfeffermann, D., and C.H. Barnard, "Some New Estimators for Small Area Means with Application to the Assessment of Farmland Values," *Journal of Business and Economic Statistics* (forthcoming). Copies of the manuscript are available from the authors.

18. Reynolds, J.E. "Analytical Studies of Factors Affecting Farm Real Estate Values," *The Farm Real Estate Market: Proceedings of a Regional Workshop, May 1984.* SNREC Publication No. 21, Mississippi State Univ., Apr. 1986.

19. Reynolds, J.E., and D. Comer. *Southeast Land Value Survey: Estimates of Agricultural Land Values for Florida, Alabama, Georgia, and South Carolina, (Final Report).* Cooperative Agreement No. 58-319V-5-00375, Univ. of Florida, Gainesville, July 1988.

20. Reynolds, J.E., and D.L. Tower. "Factors Affecting Rural Land Prices in an Urbanizing Area," *Review of Regional Studies.* Vol. 8, 1978, pp. 23-24.

21. Reynolds, J.E., and R.H. Tseng. "Factors Affecting Farmland Values Among Florida Counties," *Soil and Crop Science Society of Florida.* Vol. 29, 1969, pp. 44-48.

22. Sandrey, R.A., L.M. Arthur, R.A. Oliveira, and W.R. Wilson. "Determinants of Oregon Farmland Values: A Pooled Cross-Sectional, Time Series Analysis," *Western Journal of Agricultural Economics.* Vol. 7, 1982, pp. 211-20.

23. Scharlach, W.C., and G.E. Schuh. "The Land Market as a Link Between the Rural and Urban Sectors of the Economy," *Journal of Farm Economics.* Vol. 44, 1961, pp. 1,406-11.

24. U.S. Department of Agriculture, National Agricultural Statistics Service. *Agricultural Statistics: 1988*, 1988, pp. 377-78.

25. U.S. Department of Agriculture, Economic Research Service and National Agricultural Statistics Service. *Farm Costs and Returns Surveys, 1984-1989.*

26. U.S. Department of Agriculture, Economic Research Service. *Major Statistical Series of the U.S. Department of Agriculture: Land Values and Land Use.* Agriculture Handbook No. 671, Vol. 6, Aug. 1988.

27. U.S. Department of Agriculture, Economic Research Service. *Agricultural Land Values and Markets, Outlook and Situation Report.* AR-10, June 1988.

28. U.S. Department of Commerce, Bureau of the Census. *1982 Census of Agriculture*, 1983.

29. Wald, A. "Tests of Statistical Hypotheses Concerning Several Parameters when the Number of Observations is Large," *Transactions of the American Society.* Vol. 54, 1943, pp. 426-82.

Dynamic Specification in Econometric Estimation

James R. Malley

The use of dynamic specification with an error correction mechanism (ECM) in single-equation and multiequation macroeconomic forecasting models has become widespread over the last decade. In a single-equation context, the recent work of Hendry (4,5) on aggregate consumption and money demand has been very influential.[1] The foundations of this research rest heavily on the earlier wage-price work of Sargan (8). In a multiequation context, all the major United Kingdom macroeconomic models, including the London Business School and the Bank of England, have adopted dynamic specification and the ECM in many of their single equations. In the United States, the Federal Reserve Board's multicountry model also makes extensive use of this technique.

What is Dynamic Specification with an ECM?

Dynamic specification and the ECM do not pertain to yet another single-equation econometric estimator such as ordinary least squares *(OLS)* or minimum absolute deviations *(MAD)*. Rather, a dynamic ECM specification refers to a general class of *functional forms* that can be used to describe a particular *lag distribution*. ECM forms can be used as alternatives, for example, to Koyck or Almon lags to describe the behavior of the economy or an economic relationship as it evolves over time. As with other lag distributions, ECM forms are also used to measure the length of time with which certain target or dependent variables respond to changes in instrument or independent variables.

There are two main advantages of equations specified in this manner. First, they allow the researcher to make a distinction between shortrun dynamics and longrun equilibrium. ECM forms allow the shortrun dynamics to be determined by the data. This is desirable since economic theory usually does not offer much guidance on the shapes of lag distributions. In addition, sensible longrun equilibrium conditions can be directly imposed on the equation. Later in this note, I show how the distinction between short run and long run is made analytically and empirically.

Malley was an economist in the Agriculture and Rural Economy Division, ERS. He is currently a lecturer at the University of Stirling, Scotland.

[1]Italicized numbers in parentheses cite sources listed in the References section at the end of this article.

Second, these specifications tend to lead to forecasts that display little tendency systematically to overpredict or underpredict the variables of interest. This is because they are estimated in growth rates of the dependent and some of the independent variables, which is typically sufficient to obtain white noise residuals. The method by which certain variables are chosen to enter the equation as either levels or growth rates is not covered in this note. The interested reader should refer to the work of Hendry *(5)* for a taxonomy of the classic general-to-specific approach and to Engle and Granger *(3)* for a discussion of their two-step estimation procedure, which aims to enter variables into the regression that display the appropriate degree of integratedness.

Motivations Underlying the Development of Dynamic ECM Models

In typical economic time series data, especially annual and quarterly data, "levels" equations estimated with least squares frequently have an error structure that is not covariance stationary. In other words, the error structure is not invariant with respect to time. A nonstationary error structure is generally associated with relationships estimated by using highly trended variables. Variables containing a time trend naturally produce very high correlations with each other. Relationships among highly trended variables, however, are typically meaningless devices for describing causal links. A sufficient condition for the error structure to be stationary is that all the variables in the equation be stationary. One well-known method of attempting to deal with nonstationarity is to transform each variable from its level to its growth rate *(3)*. Growth rates are much more volatile than levels and consequently more difficult to explain using variables that have no causal link with the dependent variable.

The innovation of the ECM approach is to employ a specification that uses a combination of growth rates and levels while simultaneously attempting not to violate the basic set of assumptions in regression analysis. Equations estimated in this manner allow the relevant economic theory to enter the formulation of longrun equilibrium in levels while the shortrun dynamics of the equation are determined by growth rates. The form of a simple dynamic specification in two variables with an ECM is as follows:

$$\Delta \ln Y_t = \alpha + \beta \Delta \ln X_t + \gamma \ln(X_{t-1}/Y_{t-1}) + \varepsilon_t, \qquad (1)$$

where,

Δ = first difference operator, that is, $\Delta Y_t = Y_t - Y_{t-1}$
ln = natural logarithm
α = constant term
ε_t = stochastic disturbance term, which is jointly dependent on the distributions of X and Y.

In the following section, I will show how the simple model in equation 1 can be derived from a general autoregressive distributed lag model. This derivation is necessary to interpret the parameter estimates presented later.

Derivation of Dynamic Specification with an ECM

The following model:

$$\ln Y_t = \alpha + \beta \ln Y_{t-1} + \sigma \ln X_{t-1} + \phi \ln X_{t-1} + \varepsilon_t, \quad (2)$$

where,

α = constant term
$|\beta| < 1$
ε_t = stochastic disturbance term, which is jointly dependent on the distributions of X and Y,

can be interpreted as the modified or free-form Koyck lag model. It differs from the familiar simple Koyck lag model in that X_t and X_{t-1} are estimated freely before the imposition of the declining geometric lag distribution. The lag distribution describes the effects of X_{t-2} to X_{t-n} on Y_t and is captured by Y_{t-1}. Estimating this equation in "levels," however, can imply a somewhat limited longrun equilibrium or steady-state solution. For example, if we begin by assuming that, in equilibrium, actual Y_t equals desired Y (Y^*), then by design we will have a longrun "static-levels" interpretation to the solution. This is the assumption employed in the standard partial or stock adjustment model. Under this assumption, Y^* in the following equation is found by suppressing the disturbance term and setting actual X_t to desired X (X^*):

$$\ln Y^* = \alpha/(1-\beta) + [(\sigma+\phi)/(1-\beta)] \ln X^*, \quad (3)$$

$$\text{or } Y^* = e^{\alpha/(1-\beta)} X^{*[(\sigma+\phi)/(1-\beta)]}, \quad (4)$$

$$\text{or } Y^* = kX^{*v}, \quad (5)$$

where,

k = the constant term, that is, $k = e^{\alpha/(1-\beta)}$
v = the longrun elasticity of Y^* with respect to X^*, that is, $v = [(\sigma+\phi)/(1-\beta)]$.

The assumption that equilibrium is characterized by a single value, that is, static, underlies the derivation leading to equation 4. Sargan suggests another way to estimate equation 2 based on a more general characterization of the longrun equilibrium solution by rewriting equation 2 and estimating it in difference terms *(8)*. As described above, this procedure can generate stationary time series, which can give estimators more desirable statistical properties and the equation potentially better forecasts. This is because white noise residuals may be obtained by the appropriate order of differencing. In addition, rewriting this equation and estimating in growth rates will yield a more general longrun steady-state solution that will be constant in growth rates but dynamic in levels. Sargan's longrun steady "static-growth" but "dynamic-levels" equilibrium is derived by allowing X^* to grow at a constant rate Gx. The form of the longrun relationship between "levels" of Y^* and X^* is the same as in equation 5, that is, $Y^* = kX^{*v}$. However, the two notions of equilibrium that are used to arrive at the steady-state solutions are profoundly different. To arrive at equations 4 and 5 from equation 3, it was assumed that X would equal X^* in equilibrium. This assumption implies that the growth in X^* is zero. Sargan, on the other hand, arrives at equation 5 by assuming that, in equilibrium, the growth in X will equal the growth in X^*.[2]

Under Sargan's approach, if the steady-state growth equilibrium solution is to be maintained in the long run, then Y^* will have to grow at the rate Gy=vGx. The parameter v, by definition, is the longrun elasticity of Y^* with respect to X^*, that is, v=Gy/Gx. For example, application of the elasticity formula

$e = \partial Y/\partial X^*(X/Y)$ to equation 4 yields:

$$\begin{aligned} e &= v(Y/X^v)(X^{v-1})(X/Y) \\ &= vYX^{-v}X^{v-1}XY^{-1} \\ &= v(Y/Y)(X^{-v})(X^{v-1})X \\ &= v. \end{aligned}$$

The first step in rewriting equation 2 is to subtract the expression ($\ln Y_{t-1} + \sigma \ln Y_{t-1}$) from both sides of the equation, which yields:

$$\Delta \ln Y_t = \alpha + (\beta-1)\ln Y_{t-1} + \sigma \Delta \ln X_t + (\sigma+\phi)\ln X_{t-1} + \varepsilon_t. \quad (6)$$

---

[2]Mechanically, different assumed levels of X^* could be applied to equation 4. However, given the assumptions of the standard stock adjustment model, each newly assumed X^* implies a new steady-state equilibrium. While constant nonzero growth rates of X^* could be generated *across* equilibrium states, this is clearly not equivalent to the *within* equilibrium growth rate of X^* in the Sargan approach.

Further rearrangement gives:

$$\Delta \ln Y_t = \alpha + \sigma\Delta \ln X_t + (\beta\text{-}1)[\ln Y_{t-1} - (\sigma+\phi)/(1-\beta)\ln X_{t-1}] + \varepsilon_t. \quad (7)$$

In equation 7, the longrun elasticity (v) of Y^* with respect to X^* is $(\sigma+\phi)/(1-\beta)$. The expression for v is the same as in equation 5. Substituting the assumed longrun relationship between Y^* and X^* from equation 5 into equation 7 by setting $\ln k = \ln Y_{t-1} - v\ln X_{t-1}$, $\Delta \ln Y_t = Gy = vGx$, $\Delta \ln X_t = Gx$, and suppressing the disturbance term yields:

$$vGx = \alpha + \sigma Gx + (\beta\text{-}1)\ln k. \quad (8)$$

Further rearrangement gives:

$$k = e^{[\alpha + (\sigma-v)*Gx / (1-\beta)]}, \quad (9)$$

therefore, in the long run:

$$Y^* = e^{[\alpha + (\sigma-v)Gx / (1-\beta)]} X^{*[\sigma + \phi) / (1-\beta)]}. \quad (10)$$

Inspection of equations 5 and 10 shows that the difference between the notions of longrun equilibrium lies in the definition of k. The next feature of the Sargan and Hendry approach is the mechanics of the lag distribution as described by the error correction mechanism. The ECM term appears in equation 7, that is, $[\ln Y_{t-1} - (\sigma+\phi)/(1-\beta)\ln X_{t-1}]$. If $\Delta \ln Y_t$ starts to grow at a rate that is inconsistent with its longrun steady-state growth path, then the ECM acts to pull it back on track. $\Delta \ln Y_t$ might deviate from its longrun growth rate, for example, if there were a series of extremely large random disturbances or a systematic effect of a missing variable(s) not specified in the longrun solution. The ECM has this corrective effect because the $|\beta|<1$, making the sign of $(\beta\text{-}1)$ negative. For example, as Y_{t-1} becomes larger (smaller), the value taken by the term in square brackets become larger (smaller), but this value is then multiplied by the negative $(\beta\text{-}1)$ term, which has the effect of lowering (raising) $\Delta \ln Y_t$. A version of the ECM made popular by Hendry *(4,5)* assumes that v=1. This makes the estimation process more straightforward, since the lagged logged ratio of Y to X can be entered directly into the function as in equation 7.

At the estimation stage, it is typical to experiment with both freely estimating v and imposing the unit elastic assumption. The economic reasonableness of this assumption is judged relative to the theory underlying the relationship. Estimating v can be done quite easily by rewriting the ECM and entering $\ln Y_{t-1}$ and $\ln X_{t-1}$ directly into the relationship (see equation 6).[3]

Estimating Aggregate U.S. Consumption Expenditure

Following is a numerical example of a dynamic ECM consumption function which illustrates the concepts covered above. This equation is borrowed from the SGM model and follows the work of *(1)*. Aggregate consumption expenditure is modeled as a function of income and inflation. Consumption is defined as C, personal disposable income as Yd, and the GNP deflator as P. Using annual data from 1960 through 1988, the estimated consumption function is:

$$\Delta \ln C_t = \underset{(0.10)}{0.64}\Delta \ln Yd_t - \underset{(0.07)}{0.24}\ln(C_{t-1}/Yd_{t-1}) - \underset{(0.10)}{0.23}\Delta \ln P_t, \quad (11)$$

Mean $\Delta \ln C_t$	= 0.034
S.E. of regression	= 0.009
Adjusted R-Squared	= 0.71
Durbin Watson	= 1.76.

The standard error of each coefficient is reported underneath it in parentheses. Also, note the absence of a constant term and that v is constrained to be unit elastic. Constant terms are generally omitted in estimated difference equations since they imply a drift in the longrun steady state.

According to Hendry, the relationship estimated here is a feedback model. For example, each year consumers plan to spend the same amount as in the previous year ($\ln C_t = \ln C_{t-1}$), modified by a proportion of the annual change in income ($+0.64\Delta \ln Yd_t$) and inflation ($-0.23\Delta \ln P_t$). This determines a shortrun consumption decision, altered by $-0.24\ln(C_{t-1}/Yd_{t-1})$. The ECM term provides feedback from the previous C/Y ratio, ensuring coherence with the longrun desired outcome $C^* = kYd^*$. The implied consumption-income relation on a steady-state growth path is calculated as follows: set $\Delta \ln C_t = G1$, $\Delta \ln Yd_t = vG1$ or G1 since v=1, and $\Delta \ln P_t = G2$. Therefore, based on equations 7 and 8, we can obtain:

$$G1 = 0.64G1 - 0.24\ln C^* + 0.24\ln Yd^* - 0.23G2, \quad (12)$$

$$\text{or } \ln C^* = [(-0.36G1 - 0.23G2)/0.24] + \ln Yd^*, \quad (13)$$

[3]For example, the ECM can also be written as: $[\delta 1\ln Y_{t-1} - \delta 2\ln X_{t-1}]$, where, $\delta 1=(\beta\text{-}1)$ and $\delta 2=(\beta\text{-}1)v$. Since $\delta 1$ is uniquely determined, $v=\delta 2/\delta 1$. For pedagogic purposes and simplicity of notation, the analytic derivation of the model has been done by using a two-variable example. Extensions to the n variable model are trivial.

or $C^* = kYd^*$, (14)

where,

$k = e^{[(-0.36G1 \ - 0.23G2)/0.24]}$.

At assumed longrun 2-percent growth rates in consumption and income (G1=0.02) and 5-percent inflation (G2=0.05), the average propensity to consume (APC) or marginal propensity to consume (MPC) is k=0.93.[4] For a 0.15 G2 value, the APC or MPC falls to 0.84. From these calculations we can also solve for the longrun *steady-state elasticity* of consumption with respect to G2. As is evident from equation 13, this elasticity is a function of the assumed steady-state growth path of G1. Using the values from the original example, if G1=0.02 and G2=0.05, k=0.93. Holding G1 at a constant 2 percent and increasing G2 to a constant 15 percent (that is, a 200-percent increase), reduces k to 0.84 (that is, a 9.69-percent decline). Therefore, assuming G1=0.02, the steady-state elasticity of consumption with respect to inflation is -0.05 (that is, -9.69/200 percent). At a different assumed steady-state growth rate in G1, the elasticity value of consumption with respect to inflation will change.

Conclusions

The consumption function case was chosen because it is familiar to most economists regardless of their specialization and because it is easily adapted to more disaggregated studies. Applied economists are encouraged to consider this type of specification in their research because of the ease of application, along with the desirable statistical and forecasting properties of dynamic ECM's. Agricultural economists may be particularly interested in this example because it could be modified easily to cover consumer demand for different commodities.

[4]In the case where the longrun elasticity of consumption with respect to income is equal to 1, the APC and MPC are equal.

References

1. Davidson, J., D.F. Hendry, F. Srba, and S. Yeo. "Econometric Modelling of the Aggregate Time-Series Relationship Between Consumers' Expenditure and Income in the United Kingdom," *Economic Journal*. Vol. 88, 1978, pp. 661-92.

2. Engle, R.F., and C.W.J. Granger. "Co-integration and Error Corrections: Representation, Estimation and Testing," *Econometrica*. Vol. 55, 1987, pp. 251-76.

3. Granger, C.W.J., and P. Newbold. "Spurious Regression in Econometrics," *Journal of Econometrics*. Vol. 2, 1974, pp. 111-20.

4. Hendry, D.F. "Predictive Failure and Econometric Modelling in Macroeconomics: the Transactions Demand for Money," *Economic Modelling*. P. Omerod (ed.). London: Heinemann, 1979.

5. _____. "Econometric Modelling: the Consumption Function in Retrospect," *Scottish Journal of Political Economy*. Vol. 30, 1983, pp. 193-220.

6. _____. "Monetary Economic Myth and Econometric Reality," *Oxford Review of Economic Policy*. Vol. 1, 1985, pp. 72-84.

7. Malley, J.R., J.I. Foster, and D.N.F. Bell. "U.S. Budgetary Policy and the World Economy: the Estimation of a Small Multi-Country Model," *University of Glasgow Discussion Paper in Economics*. No. 8813, 1988, pp. 1-55.

8. Sargan, J.D. "Wages and Prices in the United Kingdom: A Study in Econometric Methodology," *Econometric Analysis in National Planning*. P.E. Hart, G. Mills, and J.K. Whitaker (eds.). London: Butterworth's, 1964, pp. 25-54.

Tweeten vs. Tweeten

Farm Policy Analysis. By Luther Tweeten. Boulder, CO: Westview Press, 1989, 399 pages, $48.50

Reviewed by Lloyd D. Teigen

A lot can change in 10 years. Tweeten's views of farm policy sure have. At least, that's my reading of the differences between his latest text and the *Foundations* he published in 1979. I'm not sure it's for the better.

Promotional material asserts that *Farm Policy Analysis* is "a completely new look at the problems of farming in the United States today," although it was originally intended as the third edition of his *Foundations of Farm Policy.* Sixteen chapters and 567 pages were reduced to 12 chapters and 399 pages in the revision. Four chapters in the *Foundations* were eliminated and four others combined into two for the current text. Two new chapters were introduced in the 1989 work.

New are chapters on "Environmental and Natural Resources" and "Poverty, Human Resources, and Rural Development." Gone are "Food and Nutrition," "Farm Organizations and Protest," "Economic History of American Agriculture: Attitudes, Institutions, and Technology," and "Economic History of American Agriculture: Interpretation and Evaluation." Chapters on goals and values for farm policy from a rural perspective and from an urban-industrial perspective are combined into a chapter on "Values, Beliefs, and Politics." "Agribusiness Conduct, Structure, and Performance" combines two chapters on market structure for purchased inputs and for farm products with new material on commodity futures markets. A glossary of agricultural policy terms is new to the 1989 book, and the appendix deriving utility measures from industry demand and supply was dropped from the 1979 text. "Public Welfare and Economic Efficiency" was renamed "Public Welfare and Economic Science."

This work deprives policy students of the context, motivation, and history of the programs affecting agriculture. I think that's a real loss. Perhaps the past is irrelevant to policy directed at removing Government from agriculture. But, even that requires knowing what values are being represented by which interest groups in order to neutralize their opposition.

Tweeten recognizes the disengagement of Government from farm policy, and economic policy generally, when he states (p. 32), "An important principle of policy analysis is that *the presence of an economic problem is not sufficient justification for public intervention* (his emphasis). It is necessary to show that the social costs of public intervention are less than the social costs of the market failures or other problems the public programs attempt to address." But, the role played by budget reform and deficit reduction legislation in this policy redirection is not discussed in his text.

Tweeten's greatest turnabout reflects his thinking on resource returns in farming. The chapter title in 1979 was "Farm Problems: Low Resource Returns," but it's "Explaining Alleged Chronic Low Returns" in the current book. His 1989 summary reads:

> Neither theory nor empirical evidence supports the hypotheses that commercial farms are chronically predestined to earn low returns in farming in the absence of government intervention. The farming industry demonstrates substantial resiliency in adapting to the uncertain environment in which it operates. This is not to deny that even well-managed commercial farms do not receive low incomes, low rates of return, and capital losses some years, and sometimes for several years. (p. 129)

Contrast that with his 1979 summary:

> Data presented in this chapter document the prevalence of low rates of return on farm resources. These are symptoms of a more basic problem of excess resources. And excess resources are a symptom of an even more basic problem--the lack of resource mobility. (p. 197)

Tweeten's 1989 conclusion has been evident for more than 40 years to analysts examining the distribution of income within the sector. The ratio of total income in the farm sector to the number of places that have or normally would have had $1,000 or móre gross sales is terribly misleading. Policy based on sectoral-average income per farm is bad policy.

Trade liberalization and market volatility illustrate other changes in Tweeten's viewpoint. Luther subscribes to Schuh's 1976 observation that greater integration of world trade and financial markets makes the U.S. farm and nonfarm economy sensitive

Teigen is an economist in the Agriculture and Trade Analysis Division, ERS.

to foreign events (p. 167). He also argues that broad-scale trade liberalization would dampen price movement throughout the world by increasing the price elasticity of demand (p. 159). In 1979, Tweeten said that free trade "could double the price elasticity of demand for U.S. exports" (p. 231), while in the 1989 text, free trade merely "would raise the price elasticity." His 1989 text criticizes the tendencies of the European Community to export its production instability, while the 1979 text said nothing about EC trade. After all, they were still buying U.S. grain at that time. Tweeten's assessment of the likelihood of trade liberalization is reflected by the elimination of:

> But in view of vested interest in protectionist policies, it would be naive to conclude that trade liberalization, however attractive to economists, will soon dampen price volatility in world markets for farm products. (1979, p. 231)

Tweeten makes widespread use of ERS descriptive and analytical reports in his discussion, and praises the agency (p. 94): "...the most comprehensive objective information from government comes from the Economic Research Service, USDA, in the executive branch. Land-grant universities, of course, also contribute in a major way." But, his sources are not consistently cited. He mentions PEST activities on p. 51 without citing Rausser as originator of the concept. He footnotes (Johnson, 1987) on p. 37 without including the work in his references. His figure 1.1 depicts commodity and multifactor terms of trade, which represent two-thirds of a figure in an uncited agricultural information bulletin on parity.

He does cite my larger report on parity in his short history of commodity programs, and defines parity prices in this glossary with better precision than the definition in the *Fact Book of U.S. Agriculture.* But, Tweeten makes a factual error when he notes that the "specific concept of parity price was used for the first time in legislation" in the Agricultural Adjustment Act of 1938 (p. 327). Parity prices were defined in the 1933 Agricultural Act. In fact, his discussion of support levels under the 1933 Act quotes corn at 60 percent of parity and cotton at 69 percent of parity (p. 325). The erroneous text did not appear in his 1979 book, and was probably intended for insertion in his discussion of the 1933 Act but escaped his editor's eye.

The additions and deletions to the opening of his chapter on "Commodity Programs: A Short History" illustrate the editorial gains and losses between the 1979 and 1989 texts. Added in the 1989 text (p. 323) was:

> A historical review and calculations of agricultural price parity are found in Teigen (1987).

Deleted from the 1979 text (p. 456) was:

> The chapter deals primarily with programs of the federal government to alleviate commercial farm problems described at some length earlier. Federal programs to alleviate problems of rural poverty are described elsewhere (Tweeten and Brinkman, 1976). An account of the excitement and drama of political infighting is sacrificed in the interest of brevity in the following pages. This is unfortunate, and the reader is encouraged to read such able presentations as *Pressures and Protest* (Hadwiger and Talbot, 1965) and *The Policy Process in American Agriculture* (Talbot and Hadwiger, 1969).

I'm pleased with his citation, but feel its wording is clumsy and not nearly as fluid as that excised from the 1979 text. Other similar examples could also be found. Generally, I found the aesthetic appearance of the 1979 text and its artwork more pleasing than that in the 1989 book.

I expected much from Luther's new text for many reasons, but I was somewhat disapppointed. The bottom line is that I would use his 1979 *Foundations*, rather than the current text, if I were teaching a course in agricultural policy. I don't think policy can be understood absent the context in which it arose or evolves. I think Tweeten erred by eliminating the contextual material.

References

Rausser, Gordon C. "Political Economic Markets: PERTs and PESTs in Food and Agriculture," *American Journal of Agricultural Economics*. Vol. 64, No. 5, Dec. 1982, pp. 821-33.

Teigen, Lloyd D. *Price Parity: An Outdated Farm Policy Tool?*. AIB-531, U.S. Dept. Agr., Econ. Res. Serv., Sept. 1987.

Tweeten, Luther G. *Foundations of Farm Policy*. Second edition. Lincoln: University of Nebraska Press, 1979.

Community Economics: A Neoclassical Synthesis

Community Economics: Economic Structure and Change in Smaller Communities. By Ron Shaffer. Ames: Iowa State University Press, 1988, 322 pages, $32.95.

Reviewed by David A. Henderson

Community economics is an abyss few academic agricultural economists have ventured into. The subject is usually subsumed in the curriculum as a subtopic in rural development, as the word "community" does not exist in a formal economics dictionary. The title "Community Economics" intrigued me and I wondered whether the book would be a reiteration of urban, regional, and developmental theories, or would the book contain new methods and insights for economists interested in the field.

The monograph contains elements of spatial theories, but I was pleasantly surprised to find a neoclassical synthesis applied to community. Although Shaffer does not break new theoretical ground, he does provide the neoclassically trained economist with the needed conceptual framework for community analysis. The neoclassical theme of the book provides a sound economic theoretical base in which community economic analysis can be framed. Shaffer adequately expands the neoclassical economic theme throughout the book and by doing so addresses the wider scope of community deveolpment. The extension of the neoclassical theme into more specific topics keeps the monograph interesting and gives it a more holistic framework.

Shaffer overcomes the lack of a precise economic concept for community by defining community from an applied perspective. Community was defined as a market area containing a political body which makes and implements decisions that affect the economic linkages in the community market. The applied framework develops the geopolitical community as a decisionmaking unit so the neoclassical efficiency principles of maxima or minima can be applied to community market problems. Market failures are introduced that violate the axioms of the neoclassical model and provide the rationale for Tinbergen-type public policy intervention at the community level.

The primary contribution of the book comes when Shaffer illuminates his intuitive knowledge of community economics. For instance, the development and discussion of multipliers in the impact analysis section of the book clarifies much of the misuse of multipliers in community analysis. The practical experience of the author is communicated with statements like:

> To a large extent the degree of leakage depends on the level of the community in the central place hierarchy....Larger communities are capable of capturing and keeping a higher proportion of spending connected with a development event (i.e., they have a larger marginal propensity to consume locally). (pp. 235-36)

A secondary contribution of the monograph is its extension orientation. Shaffer successfully takes community economics out of the academic realm and places it in the real world with his outline of community strategies and program implementation. Shaffer takes a step in the right direction by unifying the neoclassical models with the real life outreach of extension development plans.

The major weakness of the monograph is that it contains no general dynamic theory of community economic adjustment. The neoclassical equilibrium approach precludes dynamic modeling by constraining community markets to always return to an equilibrium. Shaffer completely rejects cumulative causation and stage models because of their lack of policy interjection points, but in doing so he limits community economic development to the neoclassical growth model.

The monograph is oriented toward the novice reader, but the morsels of accumulated knowledge Shaffer laces throughout his treatise will also provide insights for the rural development specialist. The rural development practitioners who have no formal training in economics will find the economic models in the manuscript useful in communicating with the professionally trained economist. The book is worth reading and should be a contribution to the reference library of those interested in rural development.

Henderson is an agricultural economist with the Agriculture and Rural Economy Divison, ERS.

The Rural Transportation Dilemma: Questions and Answers

Profitability and Mobility in Rural America: Successful Approaches to Tackling Rural Transportation Problems. *Edited by William R. Gillis, University Park, PA: The Pennsylvania State University Press, July 31, 1989, 246 pages, $24.95.*

Reviewed by T.Q. Hutchinson

"Why is there so little research in transportation?" is a question often posed by my colleagues. This book does not supply an answer but does demonstrate that the question remains valid. From its title, one would suspect that the book is a handbook for commissioners of roads. It is not. Only 1 of the 14 essays deals with a grass roots, practical issue, salting roads. Most of the others discuss the many transportation problems now confronting rural areas. Not all of these directly affect agricultural producers. Nor are the problems limited to movement of goods through marketing channels. Transportation-related problems of rural community development and viability are also shown to exist. One of the best essays deals with the lack of passenger service for the rural young and old and those physically, mentally, or economically disadvantaged. Kidder states that while deregulation removed the cross-subsidies prevalent in passenger traffic, it also removed the service. This has, to an extent, been replaced by narrowly targeted programs, which may aid the elderly but not the infirm, or the infirm but not the poor. Deregulation has produced a set of fragmented, complex programs whose efficiency is subject to question.

Most of the essays are good, some are better. I particularly recommend those by Baumel and others, Chicoine and others, and Cornelius and Kidder.

Some essays show the problems to be numerous and complex. They also show that simple solutions are bounded by constraints, with none offering more than a partial answer to a narrowly defined issue.

While some research results are presented, clearly most rural transportation questions remain unanswered. For anyone wishing to journey in a nearly unexplored land, this collection is required reading. With luck, the readers will render my colleagues' question irrelevant, and rural people and communities will benefit.

Hutchinson is an agricultural economist with the Commodity Economics Division, ERS.

The papers include: "Introduction: The Relevance of Rural Transportation" by William R. Gillis; "Extension Involvement in Rural Transportation" by Donald L. Nelson; "Alternatives for Solving the Local Rural Road and Bridge Problem" by C. Phillip Baumel, Eldo Schornhorst, and Wesley D. Smith; Financing Rural Roads and Bridges: Issues and Trends" by David L. Chicoine, Norman Walzer, and Steven C. Deller; "Deregulated Agricultural Transportation: Problems and Issues" by James C. Cornelius; "The Effect of Transportation Service on the Location of Manufacturing Plants in Nonmetropolitan and Small Metropolitan Communities" by Frank M. Goode and Steve E. Hastings; "Effective Transportation Management at the Firm Level" by James Beierlein; "Passenger Transportation Problems in Rural Areas" by Alice Kidder; "Movement of Hazardous Materials in Rural America" by Eugene R. Russell, Sr.; "Planning Rural Transportation Systems: Applications of a Statewide Network Model" by Dean Linsenmeyer, Azzaddine Azzam, and Duane Olsen; "Adjusting to a Changing Transportation Regulatory Environment—The Case of Trucking-Exempt Commodities" by Wesley W. Wilson, Gene C. Griffin, Kenneth L. Casavant, and Daniel L. Zink; "A Short-Line Solution to Rail Branch Line Abandonment: A Comparative Case Study" by Henry M. Bahn; "Highway Salt Management" by Donald J. White; "How Technology Transfer and Cooperative Extension Depasture Highways and Maintain Rural Roads" by Kate Skelton and Lynne H. Irwin.

The Book I Meant To Write

Regulation and Deregulation of the Motor Carrier Industry. *Edited by John Richard Felton and Dale G. Anderson. Ames: Iowa State University Press, 1989, 209 pages, $24. 95.*

Reviewed by T.Q. Hutchinson

This is the book that I have always intended to write. The topic is timely, the organization is logical. Individual chapters are equally logical. The writing is both clear and succinct, and it is well (while not excessively) documented.

The topic is the Motor Carrier Act of 1935, which still exists in amended form. The book can be viewed as a four-part whole: (1) Events leading to the Act of 1935. (2) The constraints stemming from the act and a major revision in 1957. These constraints included restricted entry, relatively rigid rates, and a variety of exemptions to some or all of the rules. (3) The administratively complex rules and the often inefficient industry that resulted from regulation. Especially illuminating was the point that rate bureaus restrained rate reductions, not so much by collusive or overtly coercive behavior, but by reducing the competitive advantage of lower prices. This is the best section and many insights are available. (4) A preliminary assessment of the impacts of the substantial deregulation that took place in late 1980 caps the discussion.

Many of the pieces contained in this work were published previously. Bringing them together, however, adds to their value. The book is any easy read, far easier than the task of tracking down a number of journal articles, proceedings, and monographs.

For me, the book's highlight was its focus on the lightening of economic regulation in the late 1970's through administrative fiat. Thus, deregulation has been as much a product of evolution as of legislative revoluton. The point is also made that academic opinion was a powerful force in the deregulatory movement.

While the book clearly reinforces my biases, it is the reviewer's prerogative to carp. The first section lays out the relevant events leading to regulation, but does not deal with the goals of the regulators in more than a superficial way. Based on conversations with and the writings of the academics (and others) who brought motor carrier regulation into being, I feel that a discussion of their beliefs and intentions would be useful. Without an understanding of the overall goals of regulators, it is not possible to realize the extent to which motor carrier regulation fell short.

Steeped in public utility theory, the involved academics attempted to constrain marketing of transportation services and control both inter- and intra-modal competition. In return for loss of competition's potential benefits, motor carriers were promised, nearly guaranteed, a stated return on investment. For public, single-product utilities with large startup costs, the approach remains at least moderately successful. For multiproduct transportation suppliers with nominal barriers to entry, the process was much less successful. It cannot, however, be deemed a total failure. Motor carrier regulation has lasted about 45 years and has not yet vanished.

The book also scans the influence of deregulation. This is probably not the author's fault. Deregulation brought the demise of data collection, so examining before and after conditions is difficult.

The book tantalizes with hints that deregulation has not been a complete success. Dempsey (*The Social and Economic Consequences of Deregulation, 1988)* indicates that the trend toward consolidation during 1940-72 recommenced in 1982. By 1985, the 10 largest less-than-truckload carriers enjoyed 67 percent of the market. While recourse to the Federal courts is available, such recourse is beyond the financial ability of most small shippers. Our experience with deregulated airlines has not been entirely good either. Many questions still remain.

I recommend this book to other economists. I think it contributes much, but if it does nothing else, it shows the policymaking influence of researchers for good and ill.

Australian Summary of Japan's Farm Policy

Japanese Agricultural Policies: A Time of Change. Australian Bureau of Agricultural and Resource Economics Policy Monograph No. 3, 1988, 359 pages, $44.95.

Reviewed by Lois Caplan and Larry Deaton

This work, the third in a series of policy monographs published by the Australian Bureau of Agricultural and Resource Economics (ABARE), is an accessible and valuable resource for understanding Japanese agricultural policies. It updates and is more comprehensive than a 1981 monograph by the precursor of ABARE, the Australian Bureau of Agricultural Economics (BAE). The 1988 volume puts Japanese agricultural policy into perspective by providing a good summary chapter on the development of postwar farm policies and a concluding chapter suggesting areas for further reform. The book also features separate chapters on the grains, dairy, beef, sweeteners, fruit, forestry, and fishing industries. The last two subjects were not covered in the 1981 study.

The commodity chapters are uniformly excellent. The grains chapter is as good an introduction to Japanese rice policies as we know of. Generally, these chapters not only discuss the effects of policies of a given commodity on the supply, demand, and trade of the commodity but also relate them to agricultural policies affecting other commodities. The commodity chapters also do an excellent job of providing the historical origins of these policies. For example, well known is the fact that hunger during and after World War II affected many Japanese households, having a major influence on postwar Japanese policy, particularly with regard to rice. However, it probably is not so well known, except perhaps in Japan, that major riots swept Japan much earlier, when rice prices were raised in 1918. It is probably natural for a reader to wonder why civil disturbances of 71 years ago are important to the current Japanese policy environment. The answer that ABARE gives is that the events led to the Rice Act of 1921, providing for the "permanent government control of rice and other staple foods."

ABARE goes on to weave together the story of how this law, together with more recent laws, such as the Agricultural Basic Law of 1961, have combined to solidify the structure of rice farming into its present form. Admittedly, some policy changes have occurred more recently. The ABARE book does a good job of describing the four major diversion programs, the first of which was initiated in 1971.

The other commodity chapters are equally good. The problem with these chapters is not to be found in what is there, but instead in what is left out. We were surprised that there were no chapters on poultry or oilseeds. Since Australia is not likely to be a major exporter of poultry or oilseeds to Japan, it is perhaps understandable that the book spends relatively little time on these commodities. But such an explanation does not justify the omission.

Growth in per capita poultry consumption rose from 0.8 kg./person in 1960 to 9.2 kg./person in 1985, and is likely to continue to rise. This change has had major effects on Japanese diets, significantly influencing production, consumption, and imports of other commodities, like beef.

Similarly, a discussion of oilseeds would have been useful. While formal import barriers are not apparent in the case of oilseeds imports, the domination of Japanese cooperatives in feed imports has meant that price reductions, as a result of the fall in the value of the U.S. dollar, were not passed on to Japanese livestock producers. Imports of soybean products did not accordingly increase to the extent that they would have if trade were truly free. Such omissions are not serious, but they do weaken the comprehensive nature of the work.

The chapters that do not deal with specific commodities are even better than those that do. Chapter 5, which discusses "the characteristics of Japan's agriculture," is an excellent introduction to the peculiarities of Japanese agricultural structure and the vagaries of the political environment that affect the making of Japanese agricultural policies.

Chapter 5 makes it clear that to understand the current structure of Japanese agriculture, one must understand the importance of part-time farming. Between 1960 and 1985, the percentage of farm households that derived more income from farming than from other sources declined from over 50 percent to 21 percent. Over the past three decades, the availability of off-farm income has been increasingly

Caplan and Deaton are agricultural economists with the Agriculture and Trade Analysis Division, ERS.

important to the extent that by 1986 only 15 percent of income of farm households was coming from farm sources. ABARE provides a good examination of how part-time farming has changed in recent decades, but it also goes beyond that. The book impressively goes beyond a rehash of farm structure statistics to examine part-time farming by means of a simple but illuminating labor-leisure choice model (with accompanying elasticities) in a box on page 70.

Chapter 5 also sheds light on the Japanese political process. Many writers have noted the influence of the rural voters and the debt that the ruling Liberal Democratic Party (LDP) owes to these voters. This is nothing new and the explanation here is quite straightforward. What is different, or at least less well known, is the explanation (see pp. 93-97) of the importance of the principle of noninterference by other groups in the vital interests of one group as the prime factor in allowing rural special interest groups to continue their domination of agricultural and rural policies.

The value of the information provided in chapter 5 is perhaps best understood in how it reveals the likelihood of major reforms of agricultural policies. The case for further reform is put forth in the last chapter, an especially timely and interesting piece because of the current political uncertainty in Japan. In Japanese elections of this past July, the LDP, which had been in power since 1955, lost its majority in the upper house of the Diet, or Parliament, while the Japanese Socialist Party made large gains. The LDP managed to retain its leadership in the more powerful lower house after elections were held in February. The LDP was careful not to further alienate farmers, traditional supporters of the LDP, who were partly responsible for the party's serious defeat last summer. Farmers were unhappy over the Government's lowering the rice support price in 1987 and 1988, and the impending opening of the beef and citrus markets in 1991.

As is already evident, the LDP's weakness will slow agricultural reform efforts. Just prior to the upper house elections in July, the Government, under pressure from the LDP, decided to keep the 1989 rice support price unchanged from the previous year. While the new Kaifu Government's goals for Japanese agriculture include improving productivity and ensuring a stable supply of food at reasonable prices, it has also reaffirmed its policy of maintaining self-sufficiency for rice. Many people view change in Japan's tightly controlled rice sector as key to agricultural reform in Japan.

As ABARE emphasizes, during the second half of the 1980's, the Japanese Government made several important agricultural policy changes, including lowering support prices for key commodities, such as grains, soybeans, and livestock. The Government also announced several major farm trade liberalization measures, including the removal of quotas on beef and fresh oranges. ABARE analysts view these events as starting the process of agricultural policy reform in Japan. They are also modestly optimistic that further change will occur because of both internal forces, such as the aging population of Japanese farmers, and external forces, such as U.S. pressure to open Japan's agricultural market further. The authors did not foresee, however, the political troubles of the LDP and the probable stalemate in farm reform actions. There may not exist the necessary "political will" to continue the process of agricultural reform in Japan any time soon.

One final note. It would be easy enough to restrict a review of the book to its content. As excellent as the content is, to do so would be a mistake, for the style of the work is also worthy of mention. Few publications published by a government agency of any country are as attractive as this one. How the book integrates color, text, and graphics makes it authoritative and visually impressive.

Suggestions to Contributors for Submitting Manuscripts to *The Journal of Agricultural Economics Research*

1. SOURCE. Indicate how the material submitted is related to the economic research program of the U.S. Department of Agriculture and its cooperating agencies. State your own connection with the program.
2. CLEARANCE. Obtain any approval required in your own agency or institution before sending your manuscript to one of the editors of *The Journal of Agricultural Economics Research.* Attach a copy of such approval to the manuscript.
3. ORIGINALITY OF MATERIAL. It is our policy to print original material. We consider alternative treatments of material published elsewhere, but such treatments need to differ substantially from the original approach. When submitting your manuscript, identify related material either published or submitted for publication.
4. ABSTRACT. Include an abstract and at least three keywords when you submit your article. The abstract should not exceed 100 words. It should report the major findings of your research.
5. NUMBER OF COPIES. Submit three good copies.
6. TYPING. Double space everything, including abstract, footnotes, and references.
7. FOOTNOTES. Number consecutively throughout the article.
8. REFERENCES. Check all references carefully for accuracy and completeness. Cite references in the text by using underscored numbers in parentheses that correspond to items in the reference section at the end of your article.
9. CHARTS AND OTHER ARTWORK. Use charts sparingly. Keep design as simple as possible. Submit all artwork in draft rather than in final form, accompanied by neatly prepared pages with essential data for replotting.

UNITED STATES DEPARTMENT OF AGRICULTURE
ECONOMIC RESEARCH SERVICE
1301 NEW YORK AVENUE, NW
WASHINGTON, D.C. 20005-4788

BULK RATE
POSTAGE & FEES PAID
USDA
PERMIT NO. G-145